

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 07:55:06 ; Search time 5782 Seconds
(without alignments)
3542.971 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPITRYAQTTRGLGCIIT.....PAIPDREVTYKPEDEMEEC 686

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODX=frame+ p2n model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09930591/runat_13092004_164952_700/app_query.fasta_1.839
-DB=EST -QFMT=fastap -SUFRTX=fst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09930591 @CNCN 1.1 2810 @runat_13092004_164952_700 -NCPU=6 -ICPU=3
-NOV_MMALP -LARGEBUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1:	em_estda:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estnu:*
5:	em_estcov:*
6:	em_estdpl:*
7:	em_estro:*
8:	em_hic:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_hic:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_ges_hum:*
18:	em_ges_huv:*
19:	em_ges_pln:*
20:	em_ges_vrt:*
21:	em_ges_fun:*
22:	em_ges_mam:*
23:	em_ges_mus:*
24:	em_ges_pro:*
25:	em_ges_pig:*
26:	em_ges_vrl:*
27:	em_ges_vrl:*
28:	gb_ges1:*

29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query Length	DB ID	Description
1	163	4.5	3291	11	BC046781 Mus muscu
2	150.5	4.2	3229	11	AK004665 Mus muscu
3	141	3.9	3359	29	AY411078 Homo sapi
4	140.5	3.9	822	13	BU054791 UI-M-PD0
5	136.5	3.8	3320	11	BC059369 Homo sapi
6	133.5	3.7	889	14	CD359697 AGENCOURT
7	132.5	3.7	818	13	BU187274 AGENCOURT
8	131	3.6	1201	9	AL560974
9	129	3.6	2972	11	AK031534 Mus muscu
10	129	3.6	2984	11	AK031679 Mus muscu
11	129	3.6	2986	11	AK028274 Mus muscu
12	129	3.6	3956	11	BC043699 Mus muscu
13	128.5	3.6	1697	29	AY404177 Mus muscu
14	127.5	3.5	691	13	BU054966 UI-M-PD0
15	127.5	3.5	738	14	CD240900 AGENCOURT
16	127.5	3.5	2976	29	AY400284 Homo sapi
17	126	3.5	1283	13	BO709745 AGENCOURT
18	124.5	3.4	623	14	CP131593 UI-HF-PD0
19	124.5	3.4	2388	29	AY418898 Mus muscu
20	124.5	3.4	3633	11	AY383690 Rattus no
21	124	3.4	2388	29	AY418896 Homo sapi
22	123.5	3.4	1788	11	AY105041 Zea mays
23	123	3.4	790	14	CK017540 AGENCOURT
24	123	3.4	3461	29	AY398774 Mus muscu
25	123	3.4	3802	11	AK004733 Mus muscu
26	123	3.4	4327	11	BC058331 Mus muscu
27	123	3.4	4640	11	BC062885 Mus muscu
28	122	3.4	3879	11	BC028405 Homo sapi
29	121.5	3.4	3211	11	BC041392 Homo sapi
30	121	3.3	2270	29	AY418897 Homo sapi
31	120.5	3.3	478	10	BF725559 Homo sapi
32	120.5	3.3	694	12	B1088407 Homo sapi
33	120.5	3.3	3856	11	AK078552 Mus muscu
34	120	3.3	1143	12	BM926541 AGENCOURT
35	120	3.3	3494	11	BC013208 Homo sapi
36	119	3.3	919	13	BK424517 Homo sapi
37	119	3.3	1339	14	CF753594 EST-Conti
38	119	3.3	3679	11	AK084541 Mus muscu
39	118.5	3.3	769	9	AU170397
40	118.5	3.3	1008	12	BQ050646 AGENCOURT
41	117.5	3.2	2031	11	BC038434 Homo sapi
42	117.5	3.2	3024	11	AK049321 Mus muscu
43	117	3.2	3008	11	AF318356 Homo sapi
44	117	3.2	3461	29	AY398772 Homo sapi
45	117	3.2	3915	11	BC038417 Homo sapi

ALIGNMENTS

RESULT 1
BC046781 3291 bp mRNA linear HTC 19-NOV-2003
LOCUS BC046781
DEFINITION Mus musculus DEAH (asp-glu-ala-his) box polypeptide 16, mRNA (cDNA clone IMAGE:5716182), containing frame-shift errors.
ACCESSION BC046781
VERSION BC046781.1 GI:28386173
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3291)

AUTHORS	Straussberg, R. L., Feingold, E. A., Grouse, L. H., Berge, J. G., Kleasner, R. D., Collins, P. S., Wagner, L., Shennan, C. M., Schuler, G. D., Alteschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A. A., Rudin, G. M., Hong, L., Stapleton, T. E., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. M., Brownstein, M. J., Usdin, T. B., Toshiyuki, S., Carinci, P., Prange, C., Rana, S. S., Loquellano, A. M., McEwan, P. J., Abramson, R. D., Millahy, S. J., Bosak, S. A., Loquellano, A. M., Peters, G. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worrley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulik, S. W., Villallon, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahy, J., Helton, E., Ketterman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Stevchenko, I., Bouffard, G. G., Blakeley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmitt, J., Myers, R. M., Butterfield, Y. S., Krzewinski, M. I., Skalska, U., Smalins, D. B., Scherch, A., Schein, J. E., Jones, S. J., and Marra, M. A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBLISHED	223888257		
REFERENCE	12477932		
AUTHORS	2 (phases 1 to 3291)		
TITLE	Straussberg, R.		
JOURNAL	Direct Submission		
COMMENT	Submitted (13-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
CONTACT	Contact: Gcapps-help desk		
EMAIL	Email: gcapps-remail.nih.gov		
TISSUE	Tissue Procurement: Dr. Jim Lin, University of Iowa		
CDNA	CDNA Library Preparation: M. Bento Soares, University of Iowa		
DNA	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)		
SEQUENCING	DNA Sequencing by: University of Iowa, Dr. W. Bento Soares and Dr. Thomas L. Casavant.		
WEB	Web site: http://genome.uiowa.edu		
CONTACT	Contact: bento-soares@uiowa.edu ; tom-casavant@uiowa.edu		
BONALDO	Bonaldo, M. F., Akabogu, I., Bait, T., Bait, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snit, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M. B.		
CLONE	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: http://image.ihl.gov		
SERIES	Series: Plate: Row: Column: 0		
IDENTITY	This clone was selected for full length sequencing because it passed the following selection criteria: similarly but not identity to protein		
PROBLEM	This clone has the following problem: frame shifted.		
FEATURES	Location/Qualifiers		
SOURCE	1..3291		
ORGANISM	/organism="Mus musculus"		
MOLECULE	/mol_type="mRNA"		
STRAIN	/strain="C57BL/6"		
DB_XREF	/db_xref="taxon:10090"		
CLONE	/clone="IMAGE:5716182"		
TISSUE	/tissue_type="mouse, brain 12.5 dpc"		
CLONE_LIB	/clone_lib="NIH BMAP FC0"		
LAB_HOST	/lab_host="DH10B"		
NOTE	/note="Vector: pYX-ASC"		
ALIGNMENT	Scores:		
PRED.	No.: 0.00271 Length: 3291		
SCORE	163.00 Matches: 127		
PERCENT	Similarity: 35.15% Conservative: 60		
BEST	Local Similarity: 23.87% Mismatches: 165		
QUERY	Match: 4.51% Indels: 181		
DB:	11 Gaps: 27		

US-09-930-591-2 (1-686) x BC046781 (1-3291)

QY	89	ProGlnGlyAlaArgSerLeuThrPro-----CysThrCys-----	100
Db	953	CCGGAGCGACGAGGAAGCTGGAGGCCCTAAATCGCTACCATCGCCCAAGAGACCGGAG	1012
QY	101	GlySerSer---AspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg	119
Db	1013	GCGACGACGCTCGACACTGTGTGGAATCGTTGAGAGGAATCGGATCGCCCGGAGAGAGGC	1072
QY	120	ArgGlyAspGlyArg-Gly-----SerLeuLeuSerProArgPr	132
Db	1073	AGCGGCCCTGGAGAGAGGCCACGCTGGGCTGCATCCTTGAAAGTTGGAGCCCAAGATG	1132
QY	132	oIleSerTyrLeuAspGlySerSerGlyGlyProLeuLeuCysProAlaGly-----	149
Db	1133	CTGC-----TGCACAGAGGCCAAAGTACCAAGCTGGGCTGGAGAGAGAGC	1177
QY	150	---HisAlaVal-----GlyIlePheArgAlaAl	158
Db	1178	AGACCATCGAAGTTGTCCGTGCTCTCAACTCCAGGGTGAAGAGAGCCGTAGGCCCC	1237
QY	158	avalCysThrArgGlyValAla-LysAlaValAspHeIleProValGlnSerLeuGut	178
Db	1238	CCCTGTACGCCAGACCCCAAGAGAG-----GATCATATCAAG	1276
QY	178	hrThrMetArgSer---ProValPheSerAspAsnSerSerProProAlaValProGln	197
Db	1277	CTGTGCGCGAGAGCCGTGCCGTGTTCCCTTCCGAGAGAGCTTTCGCCCGCATTTGCCA	1336
QY	197	erTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValPro	216
Db	1337	ACCATCAAGTCTCTCATCTCAAGACGACAGCTGCTCTGGAAAGACCAACAGATCCAC	1396
QY	217	-----AlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnProS	232
Db	1397	AGTACCTCTTTGAGAGAGGTTTACCAAAAGAGGCATGAAGATTTGCTTGCAACCCAGCCCC	1456
QY	232	erValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisGly-----	248
Db	1457	GGAGAGTGGCGGCTTGTGAGTGTGGAGCCCGAGTGGCCCGGAGATGGGTGTGAACCTTG	1516
QY	249	-----IleAspProAsnIleArg-----ThrGlyValArgThrIleThrT	262
Db	1517	GGAGAGAGGTGGGCTTACAGACATCCGGTTTGAGACTGACCTCAGACGAACTGTT-----	1572
QY	262	hrGlySerProIleThrTyrSerThrTyrGly-----LysPheLeuAlaAspG	278
Db	1573	-----CTCGGCTACATGACAGATGGAATGTACTCCGAGACTCTCTCTGAGC	1621
QY	278	LysGlyCysSerSerGlyGlyAlaTyrAspIleIleIleCysAspGlnCysHisSerThrAspA	298
Db	1622	CTGACCTTCA-----AGTTCAAGTGGTGATGTGGATGGAAGGCTCAAGCGGAGACT	1675
QY	298	larThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgL	318
Db	1676	TGACACACAGACATTCCTTTGGATTGATCAAAAGCTGCTTGAATTCGACCTGAGCTCA	1735
QY	318	euthrValLeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleG	338
Db	1736	AGGTCTGGTGGCTTAGCCACA-----	1758
QY	338	lueGluValAlaLeuSerThrThrGlyGlnIleProPheTyrGlyLysAlaIleProLeuG	358
Db	1759	-----CTGATATCTGCCCGGTTTTCGCTTCTCGATGAGCGCC---CTGTCT	1804
QY	358	lualIleLysGlyGlyArgHisLeu-----IlePheCysHisSerLysLysValCysA	376
Db	1805	TCAAAATCCCTGGAGCCAGGATTTCCAGTTGACATCTTATATCAAGGCCCAAGAGCTG	1864
QY	376	spGluLeuAlaAlaLysLeuValAlaAlaLeuValAlaAlaValAlaTyrTyrArgGlyL	396
Db	1865	ACTACCTGGAGGCTCGCTGTGTCTGCTG-----C	1894

QY 396 euApsValserVal1leProthrSerg1yAspValVal1Val1aThrAspAlaLeuM 416
 |||::|
 1895 TCAGATCCAGCGGACCCCGCCGAGATATACGTGG-----TTCC 1939
 QY 416 ecThg1yPheThrg1yAspPheAspSerVal1leAspCys----- 429
 |||::|
 1940 TCACGGGACA-GGAGAGATTGAGGCTGCTGAGATGCTCCAGACCGCTGCCAGG 1998
 429 ----- 429
 Db 1999 CTGGCTCCAGATCCGGAGCTCTGTGCTGCCATTATGCAACCTGCCCTCAGAC 2058
 QY 430 -----AenThrCysValThr- 434
 2059 ATGCAGGCTCGCATCTTCAGGCCACACCCCGGGGCGCGAAGAACTTCCCTCACA 2118
 QY 435 --GlnThrValAspPheSerLeuAspProThrPheThr----- 446
 |||::|
 2119 TCGAAGGCATCATCTATGTGCTGCAGCCAGGTTCTGCAAGCAGAGACCTACAACTTC 2178
 QY 447 -----llegThrThrleThrleu---ProGlnAspAlaValSerArgThrGlnArg 463
 |||::|
 2179 GTACGGGAATGAGATCTACTCAGCGTACCCCTGCAGCAGAGCTTACGCCAATCAGCGG 2238
 QY 463 rG1yArGThrg1yArG1yLysProG1y1leTyArPheValAlaProG1yGluArgP 483
 |||::|
 2239 CTGGCGCTCAGGTCA-----GTGCTGCCGG----- 2267
 Db 483 roSerGlyMetPheAspSerSerValLeuCysGlnCysTyArAspAlaG1yCysAlaTrp- 502
 |||::|
 2268 -----AAGGCTTCCGCTGTATACGCGCTGG 2295
 QY 503 --TyG1uLeuThrProAlaGluThrThrVal 512
 |||::|
 2296 CCTATCAGCATGAGTCTAGAGAGACCAAGT 2327

RESULT 2
 AK004665
 LOCUS
 DEFINITION
 AK004665 3229 bp mRNA linear hnc 20-sep-2003
 Mus musculus adult male lung cDNA, RIKEN full-length enriched
 library, clone:1200009D07 product:PROBABLE ATP-DEPENDENT HELICASE
 DD35 (DEAH-BOX PROTEIN 35) homolog [Homo sapiens], full insert
 sequence.
 ACCESSION
 AK004665.1 GI:12835999
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 99279253
 10349636

TITLE
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20499374
 11042159

REFERENCE
 AUTHORS
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichannel sequencer
 JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861

REFERENCE
 AUTHORS
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

TITLE
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

JOURNAL
 MEDLINE
 PUBMED
 6 (bases 1 to 3229)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
 Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
 Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
 Kaasakawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
 Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
 Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
 Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
 Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan [E-mail: genome-reseq@riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216]
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5' GAGAGAGAGAGCGCGCGCAACCTGAGTTTTTTTTTTTTTTN 3'], cDNA was
 prepared by using triethanolamine-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. Second
 strand cDNA was prepared with the primer adaptor of sequence [5'
 GAGAGAGAGAGATCCAGATCCAGATCAATTAATTAATTAACCCCCCCCC 3']. cDNA was
 cleaved with XhoI and SclI. Cloning sites, 5' end: SclI, 3' end:
 XhoI. Host: SOLR.

FEATURES
 source
 1. 3229
 Location/Qualifiers
 1. 3229
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM_DB:1200009D07"
 /db_xref="MG1:1904330"
 /db_xref="taxon:10090"
 /clone="1200009D07"
 /sex="male"
 /tissue_type="lung"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 1. 3229
 /note="PROBABLE ATP-DEPENDENT HELICASE DD35 (DEAH-BOX
 PROTEIN 35) homolog [Homo sapiens] (SWISSPROT Q9H521,
 evidence: FASTA, 94.2%id, 99.8%length, match=2126)"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0368 Length: 3229
 Score: 150.50 Matches: 121

Percent Similarity:	34.22%	Conservative:	72
Best Local Similarity:	21.45%	Mismatches:	221
Query Match:	4.16%	Indels:	151
DB:	11	Gaps:	24
US-09-930-591-2 (1-686) x AK004665 (1-3229)			
QY	126 SerLeuLeuSerPro--ArgProIleSerTyLeuLysGlySerSerGlyGlyProLeu	144	
DB	6 ACCATGGCTGGCCCTTGGAGACGAAATCTGGAAACCGGATACAAAGGGCCA---	62	
QY	145 LeuCySProlaGlyHISAlaValGlyIlePheAspAlaAlaValCysThrArgGlyVal	164	
DB	63 -----GGTCTACGCACTTTCTGAAAGAGCAAAAGTCTC---ACCGAAGACTCT	107	
QY	165 AlAlaValAlaAspPheIleProValGluSerLeu-----GluThrThrMetArg	181	
DB	108 GGGACAAACCATCTTATACAAATCCCTACGCTGCTCTATAGACGACGAGCGGAGAG	167	
QY	182 SerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyGlnValAla	201	
DB	168 CTGCCCGTGTCAAGCTTAGAACAACATTTGTACTGTGTAGAAAATCATACAGCCGTG	227	
QY	202 HIsLeuHISAlaProThrGlySerGlyLysSerThrLysValPro-----	216	
DB	228 GTGATTTGGAGAAACAGATGTGGAAAGAGCAACCAATTCACAGTACCTGGCAGAA	287	
QY	217 AlAlaTyAlaAlaGlnGlyTyTyLysValLeuValLeuAsnProSerValAlaAlaThr	236	
DB	288 GCTGGCTGGACAGGAGGACGAGTGTGGAGTACCCAGCTGTGCAAGTGGCTGCC	347	
QY	237 MetGlyPheGlyLalaTyMetSerLysAlaHISGlyIleAspProAsnIleArgThrGly	256	
DB	348 GTGACGATGA-TCCTTCTTCCAAAGGTTGCAAGGCCAGTACGTATGAAGGGGTGA	406	
QY	257 ValArgThrIleThrThrGlySerProIleThrTy-----Ser	269	
DB	407 GTCTGGGCGCATGAAGGGCTACTGTATCGGCTGTGATGATGACCCGACCACTGGCC	466	
QY	270 ThrTyGlyLysPheLeuAlaAspGlyGlyCysSerGly-----	282	
DB	467 ACCAGATCAAGTTCCTTACAGATGGCATGCTGTGACAGAAATGATGGTCCGCTG	526	
QY	283 ---GlyAlaTyAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSerIle	301	
DB	527 TTAAACAAATATATGTCATCATCTGTGATGAAGCCACAGAGACCTTGTACAGGAC	586	
QY	302 LeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeu	321	
DB	587 ATGGCCATGGCTGTGAGAGAAAGATTCAGAAAAGCGAGGAGATCTTCGCTGATTTGTG	646	
QY	322 AlAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGlu----	339	
DB	647 GCCTCAGCCACTGTGACGACGAGAAATTCGAGATTTCTTTAACCAAGATGAGACCACT	706	
QY	340 -----ValAlaLeuSerThrThrGlyGlu-----	347	
DB	707 GACCCACGACAGATACCTCTGTGACGCTCAGGTGAGAGACGACATTTCCAGTGGAT	766	
QY	348 -----IleProPheTyGlyLysAlaIleProLeuGluAlaIle	360	
DB	767 ATCTTTTACCTACAAAGTCCGTTCGAGATTATATCAAGCAACTGTGACACCGTGTGA	826	
QY	361 Lys-----GlyGlyArgHisIleLeuPheCysHisSerLysValLys	374	
DB	827 AAAATTCATCAGACAGAGAGATGAGACATCTAACCTTTCTTACTGGCCAGAGAA	886	
QY	375 CysAspGluLeuAlaAlaLysLeuVal-----AlaLeuGly-----	386	
DB	887 GTAAGACATGTTGTGCATGTATGATGAGACAGCCCGGGGCTGGCTCGCACTGGATG	946	
QY	387 -----ValAsnAlaValAlaIleTyTyArgGlyLeuAsp-----	397	

DB	947 AAGAAACACTCCGGGTCTCCCATGATGATGACAGACTGCCCTTTTGAACAGATGAAG	1006	
QY	398 ---ValSerValIleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMet	416	
DB	1007 GTGTTTGAAGAGGTGCACACAGAGCTCAGAAAGTATCGTGGCCACCAATGTGGCAGAG	1066	
QY	417 ThrGlyPheThr---GlyAspPheAspSerValIleAspCysAsnThrCysValIleGln	435	
DB	1067 ACTTCATCACAATCAGAGGATGTGTGATGTGATGTGACTGTGGCTTTATGAACTGGCA	1126	
QY	436 ThrValAspPheSerLeuAspProThrPheThrIleGlu---ThrIleThrLeuProGln	454	
DB	1127 -----GCTTACAAACCCAGACAGCACTATGTATGATCTGTGGTGTGACCACTG	1174	
QY	455 AspAlaValSerArgThrGlnArgArgGlyValArgThrGlyArgGlyLysProGly---Ile	473	
DB	1175 TCTCAGGGCTGACCAATCAGCGGGCAGAGTGTGGGGCCAAACCGCTCGGAAAGTGT	1234	
QY	474 TyArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSerValLeuCys	493	
DB	1235 TATGCTCTTACACA-----GAGGAAGCTTTTGGACAGCTACCTCAGTCCAGCTCCT	1288	
QY	494 GluCySerTyArgPheAlaGlyCysAlaIleTyArgGluLeuThrProAlaGluThrThrValArg	513	
DB	1289 GAGATGCAAGCGACCAATTTG-----GCCCGCTGATCCTGACAG	1327	
QY	514 LeuArgAlaTyMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPheThr	533	
DB	1328 CTMAAAGC-----	1336	
QY	534 GluGlyValPheThrGlyLeuThrHisIle---AspAlaHisPheLeuSerGlnThrLys	552	
DB	1337 -----CTAGGATACACATAGTCTCCAGGTTCACCTTCATAGTCT-----	1375	
QY	553 GlnSerGlyGluAsnLeuProTyLeuValAlaTyGlnAlaThrValCysAlaArgAla	572	
DB	1375 -----	1375	
QY	573 GlnAlaProProSerThrAspIleMetTyLysCysLeuIleArgLeuLysProThr	592	
DB	1376 -----CCCCGCCAGCA---CAGTGCATGGTTCAGGCTTG-----	1408	
QY	593 LeuHisGlyProThrProLeuLeuTyArgArgLeuGlyAlaValGlnAsnGluValThrLeu	612	
DB	1409 -----GAGCTGCTTATGCTCTCGAGGCTGTGACAAAGACTGTGGCTTA	1453	
QY	613 ThrHisProVal	616	
DB	1454 ACTGAGCTCTT	1465	
RESULT 3			
AY411078 3359 bp DNA linear GSS 16-DEC-2003			
LOCUS Homo sapiens TIE gene, VIRUAL TRANSCRIPT, partial sequence,			
DEFINITION genomic survey sequence.			
ACCESSION AY411078			
VERSION AY411078.1 GI:39767046			
KEYWORDS GSS.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
REFERENCE 1 (bases 1 to 3359)			
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,			
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,			
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,			
Adams,M.D. and Cargill,M.			
Inferring nonneutral evolution from human-chimp-mouse orthologous			
gene trios			
JOURNAL Science 302 (5652), 1960-1963 (2003)			
PUBMED 14671302			
REFERENCE 2 (bases 1 to 3359)			
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,			
AUTHORS			

Todd, M.A., Tanenbaum, D.M., Civejlo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES
Location/Qualifiers
1..3359

source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3359
/gene="TIE"
/locus_tag="HCM4104"

ORIGIN

Alignment Scores:

Pred. No.:	0.323	Length:	3359
Score:	141.00	Matches:	168
Percent Similarity:	31.38%	Conservative:	84
Best Local Similarity:	20.92%	Mismatches:	233
Query Match:	3.90%	Indels:	318
DB:	29	Gaps:	41

US-09-930-591-2 (1-686) x AY411078 (1-3359)

```

QY 39 ThrAlAlGlnThrPheLeuAlaThrCysGlyLeuArglyValCysTrpThrValTyrlHis 58
DB 42 ACGGACCCCGGAGGCTTCTTCGACTGCGTGTGGGAGGCGC----- 86
QY 59 GlyAlaGly----- 63
DB 87 GGGGGGGGAGGGGCTCGACGCTGGGCGCGCCCTGCTGTGAGAAAGACGACCGT 146
QY 64 ThrLeuAlaSerProlysglyProvalIleGlnMetTyrlThrAnValAspGlnAsp--- 82
DB 147 ATGTGCGGACCCCGCGCGGCGACCCGCTGCGCTGGCGGACGAGTTGCGACAGGTC 206
QY 83 --LeuValGlyTrp-ProAlaPro----- 89
DB 207 ACGCTTCGGGCTTCTTCGACGCTCGACCTGCGGCGCTTCTTCGCTGCGGCGGT 266
QY 90 -----GlnGlyAlaArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrl 107
DB 267 GCTGGGGCGGGCGGCGGCGGCTGATCTACGTGACAAAGCCCTGAGCCCACTGCTT 326
QY 107 uValThrArgHis-----AlaAspValIleProVal----- 117
DB 327 CCGAGACAGGTGCACACACTGTGAACAAAGGTGACACGCGCTGACTTCTGACGTTG 386
QY 118 -----ArgArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProIle 133
DB 387 CACAAAGAGAAAGACAGACAGCTGATCTGGAAAGACAGACGATCTCACTTCAACCTG 446
QY 133 eSerTyrlLeuTyrglySerSerGlyTyrlProLeuLeuCysProAlaGlyHisAlaValG 153
DB 447 GACTGGCATGAAAGCCAGATGGGCGGTTCC-----TGCTGCAAGTCCCAAAATGGCGAG 500
QY 153 yIlePheArgAlaAlaValCysThrArgGlyValAlaValAlaAspPheIleProVal 173
DB 501 CCAACATCGAGCGGCA-----TCATCAAGTGCACACT 530
QY 173 lGlnSerLeuGlnThrThrMetArgSerProValPheSerAspAsn----- 188
DB 531 TACCTGAAAGCCAGCCCTGCGGACAGCGCTTCTTTCGCTCATCTGCGGCGGTTGGG 590
QY 189 -----SerSerProAlaValProGlnSerTyrlGlnValAlaHisLeu----- 203
DB 591 GCTGGCGCTGGGGCGGCGGCGGCTGTACCAAGAGTCCCAAGTTGACTACAGAGGTTTC 650
QY 204 -----HisAlaProThrGlySer---GlyLysSerThrTyrl 214

```

```

DB 651 TCCCAAGACCATGAGCGGAGATGTGTATGCCCCCTGCTTCACTGGCAACCCGCTGTGA 710
QY 214 eValProAlaAlaTyrlAla----- 225
DB 711 CAGGCTTCGAGAGAGGCGCGTTTGGGACAGCTGCCAGAGACAGTCCAGCATATCA 770
QY 225 eValLeuValLeuAsnProSer-----ValAlaAlaThrMetGlyPheG 240
DB 771 GAGTCGCGGGCGCTCATCTTCTGCTCCCAAGCCCTATGGCTGTGCTCTTGATGTGCG 830
QY 240 yAlaTyrlMetSerTyrlAlaHisGlyTyrlAspProAsnIleArgGlyValArgThrI 260
DB 831 TGGAGAGGAGCGGAGCCCAAGACCTTGCTCCCTGTCATTTGGG----- 879
QY 260 eThrThrGlySerProIleThrTyrlSerThrTyrlGlyLysPheLeuAlaAspGlyCys 280
DB 879 ----- 879
QY 280 eSerGlyAlaTyrlAspIleIleCysAspGlyCysHisSerThrAspAlaThrSe 300
DB 880 -----CTGATTCGCGACTCCAGTGCAGTGCAGATGTGCAATGTGCG 917
QY 300 rIleLeuGlyIleGlyThrValLeuAspGlnIleGlnThrAlaGlyAlaArgLeuTyrl 320
DB 918 ACTTGTG--ACCGGTTCAAGTGTGTGTCTGCTCCCTCGGGGTGCATGAGTGCACGT 974
QY 320 lLeuAla---ThrAlaThrPro----- 326
DB 975 GAAAGTACAGCGGATGCCCAAGTCTCAACATGCGCTCAAGATGCACTTCACTTA 1034
QY 327 -----ProGlySerValThr-----ValProHisPro----- 335
DB 1035 GAGAGATGCCCGGATCACTGTGACGTGCAGGAAACCTTCCCGGCGGGCGAGC 1094
QY 336 -----AsnIleGlnGlyAlaAlaLeuSerThrThrGlyGlnIleProPheTyrlGly 353
DB 1095 ATAGAGCTACGAGACAGACGCGCACTGTCTCTGCCACCAAGGCCATTGTGAGGCA 1154
QY 353 sAlaIlePro-----LeuGlnAlaIleLeuGly 363
DB 1155 GAGAAAGCCACAGCTGAGTTCAGAGTCCCGCTGTTCTTGGGACAGTGGCTTCTG 1214
QY 363 yArgHisLeuIlePheCysHisSerTyrlValLysCysAspGlnLeuAlaIleLeuVal 383
DB 1215 GAGTGGCGG-----TGTCCACATCTGGCGGCAAGACAGC 1250
QY 383 l-AlaLeu-GlyValAsnAlaValAlaTyrlTyrlArgGlyLeuAspValSerValIlePro 402
DB 1251 CGGCGCTCAAGTCAAT-----GTGAAGTGCCTCC 1283
QY 403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThr----- 417
DB 1284 -----GTGCCCTGTGCTGCACTCGGCTTCTTCAACAGAGACGCGCCAG 1328
QY 418 -----GlyPheThrGlyAsp----- 422
DB 1329 CTGTGTGTTCCCGCTGCTGCTTCTGTGGAGTGAACCATCTCCACTGCGGCTG 1388
QY 423 -----PheAspSerValIleAspCysAsnThrCysValThrGlnThrValAsp 438
DB 1389 CACTACCGGCGCCAGAGACATGACATGACTGTGCAACATTTGTG----- 1433
QY 439 PheSerLeuAspProThrThrPheThrIleGlnThrIleThrLeu-----ProGln 454
DB 1434 -----GTGGAGCCAGT-----GAGAAAGTGAAGCTTAAATGAACCTGAGGCCAAAG 1478
QY 455 AspAlaValSerArgThrGlnArgArgGlyArgGlnGlyArgGlyLysProGlyIleTyrl 474
DB 1479 ACAAGATACAGTGTCTGTGTGACGTGAGCCGCGGAGGAGGAG----- 1523
QY 475 ArgPheValAlaProGlyGln-ArgProSerGlyMetPheAspSer-----SerVal 491

```

```

Db      1524 -----GGAGAGGGGGCCCTGGGGCCCTCCACCCCTCATGACACAGACTGT 1568
Qy      491 lleucysglucyrytaspalaglycysala--trpyrgylleuthrproalagluThr 510
      1569 CCTGAGCCCTTGT-----TGCAGCGTGGTGTGAGGGCTGCATGTGGAAGGC 1616
Qy      511 ThrValArgLeuArgAlaTyrmetaenThrProGlyLeu-----ProValCysGlnAsp 528
      1617 ACTGACCGGCTGTGAGTGTGAGTGTGCTTGTGCTGTGTCGCGGCGCACTGTGGGCGAC 1676
Qy      529 His-----LeuGluPheTrpGluGly----- 535
      1677 GGTTCCTGTGCTGGCTGTGGAGCGGACACACGGGGCGAGAGCGGGAACGTCTCA 1736
Qy      536 -----ValPheTrpGlyLeuThrHisIleAspAlaHisPheLeu 548
      1737 TCCCCCGAGCGCCGACCTGCCCTCTGACGGGACTCAGCCCT--GGACCCACCTACACAG 1793
Qy      549 SerGlnThrIleGlnSerIleGlyGluAenLeuProTyLeuValAlaTyrglnAlaThrVal 568
      1794 CTGGATGTGTGAG-----CTTACCACTGCACCTTC 1823
Qy      569 CysAlaArgAlaGlnAlaPro-----ProProSerTrpAspGlnMetTrp 583
      1824 CTGGGCGCGCGCTCGCGCCCTGCACACGTGCTTCTGCCCCCGAGTGGCT----- 1874
Qy      584 LysCysLeuIleArgLeuLysProThrIleHisGlyProThrProLeu---LeuTyArg 602
      1875 -----CCAGCCCCCGACACCTTCACGCC 1898
Qy      603 LeuGlyAlaValGlnAenGlnValThrLeuThr-----His 614
      1899 CAGGCGCTCTCAGACTCCGAGATCCAGCTGACATGAGAAGACCGGAGGCTGTGCTGG 1958
Qy      615 ProValThrIleTyrlleMetThrCysMetSerAla----- 626
      1959 CCATATCCAGATGACGTGTGTGAGGTGACAGTGTGGGGTGACAGAGACCCACGTGTGG 2018
Qy      627 ---AspLeuGlnValValThrSerThrTrpValLeuValGlyValLeuAlaIa 644
      2019 ATAGACGTGTGACAGGCGCTGTGAGAGACAGACCATCATCTCGTGTGCTCAACGCCAG 2075

RESULT 4
BU054791      822 bp      mRNA      linear      EST 26-AUG-2002
LOCUS      IMAGE:6404900 5', mRNA sequence.
DEFINITION      BU054791
ACCESSION      BU054791
VERSION      BU054791.1 GI:22494868
KEYWORDS
SOURCE
ORGANISM      Mus musculus (house mouse)
COMMENT
  Mus musculus
  Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 822)
  NIH-MGC http://mgs.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cga@bbs-remail.nih.gov
  Tissue Procurement: Dr. James Lin, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LINL at:
  http://image.lnl.gov
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP)
  Seq primer: PYX-5.
  Location/Qualifiers
    1..822
    /organism="Mus musculus"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      0.0303      Length:      822
Score:          140.50      Matches:      73
Percent Similarity: 35.33%      Conservative: 45
Best Local Similarity: 21.86%      Mismatches: 129
Query Match:      3.88%      Indels:      87
DB:              13      Gaps:      11

```

US-09-930-591-2 (1-686) x BU054791 (1-822)

```

/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6404900"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NIH BMAP FDO"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is TAGAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

```

```

Qy      98 CysThrCysGlySerSerAspLeuThrArgHisAlaAspValIleProVal 117
      12 TGTGCTGGCGGAGGAGGAGCAAGTCAAGTGTGTGAGGAGGAGAGACCATCAGTTT 71
Qy      118 ArgArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyLeuLys 137
      72 GTCCGCGTGTCTCAATCCAGGGTGAC-----GAG 101
Qy      138 GlySerSerGlyGlyProLeuLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAla 157
      102 GAGCCGTCAGGCGCGCCCTGTACGCCAGCGCCAGCAGAG----- 143
Qy      158 AlaValCysThrArgGlyValAlaValAlaValAspPheIleProValGlnSerLeuGlu 177
      144 -----GAATCTATCCAG 155
Qy      178 ThrThrMetArgSer---ProValPheSerAspAsnSerSerProProAlaValProGln 196
      156 GCTGTGCGCGGAGAGCTGTGCGCGGTTCCTTCGAGAGAGACTTGTGGCCCATTTGCC 215
Qy      197 SerTyrglnValAlaHisIleHisAlaProThrGlySerGlyLysSerThrTyValPro 216
      216 AACCATGAGTCTCTCATCTCAAGGCGGAGACTGTGCTGGGAAGACCAACAGATCCCA 275
Qy      217 -----AlaAlaTyraAlaAlaGlnGlyTyrlleValleuValleuAsnPro 231
      276 CAGTACCTCTTGGAGAGAGGTTACACAAAGAGGCGCATGAAATGCTTGCACCCAGCCC 335
Qy      232 SerValAlaAlaThrMetGlyPheGlyAlaTyrmeserTyAlaHisGly----- 248
      336 CGGAGAGTGGCGGCTGTGAGTGTGCGAGCCGAGAGTGGCCGGGAGATGGGTGTGAAGTT 395
Qy      249 -----IleAspProAsnIleArg-----ThrGlyValAlaThrIleThr 261
      396 GGGAGACAGGTGGGCTACAGCATCCGTTTGGAGACTGCACCTCAAGGCAACTGTT--- 452
Qy      262 ThrGlySerProIleThrTySerThrTyrgly-----LysPheLeuAlaAsp 277
      ::: ||| ||| |||

```

D	b		453	-----CTCCGGTACATGACAAGTAGAATGCATCCGAGGTTCCTCTCGAG	500
O	y		278	GtAglyCySsergylYAlAtyrAspIlellelCYsAbpGlucyHisserThrasp	297
D	b		501	CCTGACCTTGCA-----AGTTACAGTGTGGTGAATGGTGAAGAAGCTCAGAGCGCAC	554
O	y		298	AlAthrserlLeuglYlIgLythrValIeuAspGlnalagUThrAlaglyAlary	317
D	b		555	TTCACACAGACATTCCTCTTTGGATTGATCAAAGACGTGCGTAAGATTCGACCTGAGCTC	614
O	y		318	IeuthrValIeuAlaThrAlathrProProglYserValThrValProHisProAsnile	337
D	b		615	AAGTCCCTGGTGGCTTCAGCCACA-----	638
O	y		338	GlugluValAlaIeuserThrrhgylulIeProPheTyrglyLYalaIlleProIeu	357
D	b		639	-----CTGGATACACTGCCCGTTTTCTGCTGCTTCGATGAGACCCTGCTTCA	686
O	y		358	GlualAllelyBglyYlArghIsIeu--llPheCyHisserlyslYsbysCyAsap	377
D	b		687	GAATNCNCGACGACGAGGTTTCAGTGTGACATCTTTATACAAAGGCCCCAGAGCGTACT	746
O	y		377	IuIeuAlAlalyseuValAlaleuglyValAsnAlavalAlatyrTrArglyLyeu	397
D	b		747	ACCCTGGAGCCTGCTGTGTCTGTG-----CTGC	776
O	y		397	spvalSerValIleProThrSerghYaspaValval	409
D	b		777	AGATTCACTGACCCAGCCCTCGAGATATATCTGTG	814
RESUT_5					
BC059369	LOCUS				
DEFINITION				Homo sapiens cDNA clone IMAGE:30346256, containing frame-shift errors.	
ACCESSION				BC059369	
VERSION				BC059369.1	
KEYWORDS				GI:37590735	
SOURCE				HTC.	
ORGANISM				Homo sapiens (human)	
REFERENCE					
AUTHORS					
				1 (bases 1 to 3320)	
				Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klisner,S.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Bulow,K.H., Schnefer,C.F., Bat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Dietchenko,L., Marzine,K., Farmer,A.A., Rodin,G.M., Hong,L., Stepleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Udell,T.B., Toshylycki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mollaby,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.W., Gay,L.J., Huily,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywnicki,M.I., Skalska,U., Smalls,D.E., Schenker,A., Schein,J.E., Jones,S.J., and Marra,M.A.	
				Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
				Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
JOURNAL MEDLINE PUBLISHED				12478937	
REFERENCES				2 (bases 1 to 3320)	
AUTHORS				Strausberg,R.	
JOURNAL TITLE				Direct Submissions	
JOURNAL				Submitted (01-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cga@bs-remail.nih.gov

Tissue Procurement: Dr. Stefan Hansson

cDNA Library Preparation: Michael Brownstein / Ted Ustin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-sbgc.stanford.edu>

Contact: (Dickeson, Mark) mcd@pxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at: <http://image.lnl.gov>

Series: IRAX Plate: 133 Row: 1 Column: 8

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 20544128

This clone has the following problem: frame shifted.

Location/Qualifiers

source	1..3330
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="IMAGE:30346256"	
/tissue_type="Placenta, normal"	
/clone_id="NH MGC_147"	
/lab_host="DH10B"	
/note="Vector: pBluescript"	

ORIGIN

Alignment Scores:

Pred. No.:	0.836	Length:	3320
Score:	136.50	Matches:	114
Percent Similarity:	33.6%	Conservative:	73
Best Local Similarity:	20.54%	Mismatches:	216
Query Match:	3.77%	Indels:	153
DB:	11	Gaps:	24

US-09-930-591-2 (1-686) x BC059369 (1-3320)

```

QY      132   ProIleSeTyreuleuVsglySerSerGlyProleuleuCySproalaglyHhSaLa 151
          |||:::  ::::  ::::  ::::  |||||
Db       35   CCGGTGAAGTTCCTGGACCCGGACAAGAAGGCCA-----GGTGTAAC  79
QY      152   ValGlyIlePhearagaalaalavalCysThrArgGlyValAlaYsaIaValaPheile 171
          ::::  ::::::  ::::::  ::::::  |||
Db       80   ATCTCTGAAGAGAACAAAGCTGGCTGAACACTCT---GGACAAACGTTGTTTAAC  136
QY      172   ProValGluseRleu-----GLuThrThMecharGerProValPheSerAspAn 188
          |||  ::|||  ::::  ::::  ::|||
Db      137   CCTTAGCTGCCCCCTTTCATAGACACAGACAGACAGAACGTCGCCGATTCAAGCTTAGG 196
QY      189   SerSerProProalavalProginSerTyGlnValAlahisIseuhIsalaProthrGly 208
          :::  ::::  ::::::  ::::::  ::::
Db     197   AATCATATTTTATACCTGATAGAAAATTATCACAGCGGTGTGATTGTTGTTGAACAGGA 256
QY      209   SerGlyLySerThrylsValPro-----AlaIatYrzlaIaIagIngly 223
          |||
Db     257   TGTGGGAAGGACACACAGATTCCTCACTAGTACTTCGACAAACCAGCCGCTGACAGCTGAAGGA 316
QY      224   TyrltyrValIeulleuAnProserValAlaIathMetGlyPheglValarymet 243
          |||  |||  |||  ::::  ::::
Db     317   AGAETGTAGAGTAGACCCAGCCTCCGAAGAGTGCGCTGCTGTACAGTTCGACGGAGAAGTA 376
QY      244   SerIySaIahIsgIylIeaepProasnhIeargThrglyValArgrThriIethrthrgly 263
          ::::  ::::  |||  |||  ::::
Db     377   GCTGAAGAAAGGGGTGC-AAGTGTGGGCCACAGAGTGGGCTACTG----- 420
QY      264   SerProIle-----ThTrYserThTrTyrltylYsPheleulaIaapgly 278
          ::::  ::::  ::::  ::::
Db     421   CATCCGCTGTAGTACGACGACGACGACGCTGACGAGAAATTAAATTTCTTACTGTATGGA 480
    
```



```

Qy 218 AlaTyrAlaAlaGlnGlyTyrLeuValLeuValLeuAsnProSerValAlaAlaThrMet 237
Db 164 GGTATATCAACAAGGATATGAAGATTGCTGCACCAACCCGAGAGTGGCTGCATG 223
Qy 238 GlyPheGlyAlaTyrMetSerLeuAlaHisGly-----IleAspPro 251
Db 224 AGTGTGGCCGCCGAGTGGCCGAGAGTGGTGAAGCTTGGAAATGAGTTGGCTAC 283
Qy 252 AsnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleThr 267
Db 284 AGCATCGGCTTTGAGAGCTGCACATCAGAGCAACTGTC-----CTCCGC 328
Qy 268 TyrSerThrTyrGly-----LysPheLeuAlaAspGlyGlyCysSerGlyGly 283
Db 329 TACATGACAGATGGATGCTTCTCCGGAGTTCCTCTCTGAGCTGACCTGGCG----- 382
Qy 284 AlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSerIleLeuGly 303
Db 383 AGTTACAGCGCTGGATGGATGGATGAGCACAAGAGACCTTACACAGACATTTCTC 442
Qy 304 IleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThr 323
Db 443 TTGGATGATGATCAAGATGTTGCTCGCTCCGAGCTGAGCTCAAGGCTCGTGGCTTCA 502
Qy 324 AlaThrProGlySerValThrValProHisProAsnIleGluGluValAlaLeuSer 343
Db 503 GCCACA-----ATGCACTGCGCCGCTTTTTC 529
Qy 344 ThrThrGlyGluLeuProPheTyrGlyValAlaIleProLeuGluAlaIleLysGlyGly 363
Db 530 ACC-----TTCTTTGATGAGCGCC---CCTGTGTTTCGAATCCCGGACGC 571
Qy 364 ArgHisLeu-----IlePheCysHisSerIleLysCysAspGlyLeuAlaAlaLys 381
Db 572 AGCTTTCTCTGACATCTTCTACACCAAGGCTCCAGAGCTGACTCTTGGAAAGCTTGT 631
Qy 382 LeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIle 401
Db 632 GTAGTATCTGTG-----TTGAGATTCAGATGTACC 661
Qy 402 ProThrSerGlyAspValValValAlaThr 412
Db 662 CAGCCCGCTGGGATATCTGTGTTCTGTGACA 694

RESULT 7
BU187274 818 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7825134 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6142792
5', mRNA sequence.
VERSION BU187274.1 GI:22701258
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 818)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LAM3465 row: C column: 17
High quality sequence, stop: 546.
Location/Qualifiers
1..818
FEATURES
Source

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6142792"
/cisue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site: 1; Note:
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN

Alignment Scores:
Pred. No.: 0.169 Length: 818
Score: 132.50 Matches: 63
Percent Similarity: 41.67% Conservative: 37
Best Local Similarity: 26.25% Mismatches: 94
Query Match: 3.66% Indels: 46
DB: 13 Gaps: 13

US-09-930-591-2 (1-686) x BU187274 (1-818)

Qy 128 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 147
Db 50 CTCTCAGAGGCCCAAGATCACTGCTGCTGAGAGAGAGACCAATTGAGTTGTCCG 109
Qy 148 AlaGlyHisAlaVal-----GlyIlePheArgAlaAla-ValCysThrArgAl 163
Db 110 ---GGCCACTCAGCTCCAGGCTATGAGAGCCGTCAGCTCCACCCACTTCAATCAGGC 166
Qy 163 ValAlaValAlaValAspPheIleProValGluSerLeuGluThrThrMetArgSer-- 182
Db 167 CCAGCAGAA-----GAGTCATCAGAGCCGCTCCGCGCAGCCT 205
Qy 183 -ProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnValAlaH 202
Db 206 CCGGCTTCCCATTTGAGAGAGAGCTCTGCTGCTATTGCAATCAACCAAGTCTCAT 265
Qy 202 IleuHisAlaProThrGlySerGlyLysSerThrLysValPro-----Al 217
Db 266 CATTGAAGGCGAGACAGCTCAGAGAAAGACACCGCATCCGAGTATCTTTGAGCA 325
Qy 217 AlaTyrAlaAlaGlnGlyTyrLeuValLeuValLeuAsnProSerValAlaAlaThrMe 237
Db 326 GGGTATATCAACAAGGATATGAAGATTGCTGCACCAACCCGAGAGTGGCTGCAT 385
Qy 237 GlyPheGlyAlaTyrMetSerLeuAlaHisGly-----IleAspPr 251
Db 386 GAGTGTGGCCGCCGAGTGGCCGAGATGGGTGTAAGCTTGGGAATGAGGTTGGCTA 445
Qy 251 AsnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleTh 267
Db 446 CAGCATTCGCTTTGAGAGACTGCACATCAGAGCAACTGTC-----CTCCG 490
Qy 267 TyrSerThrTyrGly-----LysPheLeuAlaAspGlyGlyCysSerGlyGly 283
Db 491 CTACATGACAGATGGATGCTTCTCCGAGATTCCTCTGACACCTGACCTGGCG----- 545
Qy 283 ValaTyrAspIleIleCysAspGlyCysHisSerThrAspAlaThrSerIleLeuGly 303
Db 546 -ACTTACAGCGTGGATGGATGGATGAGCACAAGAGACCTTACACACAGACATTTCT 604
Qy 303 YllleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThr 323
Db 605 CTTTGATTTGATCAAGATGTTCTCGCTTCCGA---CTGAGCTCAAGGCTCTGGAGTGG 661
Qy 323 AlaThrProArgProGlySerValThrValPro-----HisProAsnIleGluGly 339
Db 662 CTTCAGCACAATGGG-----ACACTGCCCGCTTTTTCACACCTTCTTTATGATA 713

RESULT 8
AL560974

```

LOCUS AL560974 1201 bp mRNA linear EST 31-MAY-2003
 DEFINITION AL560974 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0DL006YA04 5-PRIME, mRNA sequence.
 ACCESSION AL560974
 VERSION AL560974.2 GI:31285103
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 REFERENCE 1 J.M.B., Gruber, C., Jesse, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12907950.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3796.r For more information about this cluster, see
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DL006BA02QPI&cluster=3796.r. Contact: Peng Liang Email: filiang@life.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID: CS0DL006BA02QPI.
 FEATURES
 source
 1. 1201
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL006YA04"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime and enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.458 Length: 1201
 Score: 131.00 Matches: 72
 Percent Similarity: 37.11% Conservative: 46
 Best Local Similarity: 22.64% Mismatches: 109
 Query Match: 3.62% Indels: 92
 DB: 9 Gaps: 14
 US-09-930-591-2 (1-686) x AL560974 (1-1201)
 QY 128 leuSerProArgProIleSerTyLeuylsGlySerSerGlyGlyProleuLeuCyPro 147
 DB 123 CTCTCAGAGGCCCACTGATCACTGCTGCTGAGAGGAGGAGACCATGAGTTGTCG 182
 QY 148 AlaGlyHisAlaVal-----GlyIlePheArgAlaAla-ValCysThrArgG1 163
 DB 163 ---GGCCACTCAGCTCCAGGGGTGAGAGGAGCGGTCACTCCACCATTCATCACTGAGC 239
 QY 163 yValAlaIysAlaValAspPheIleProValGluSerLeuGluThrThrMetArgSer-- 182
 DB 240 CCAGCAGAAA-----GAGTCCATCCAGCGCGTCCGCCGACGCT 278
 QY 183 -ProValPheSerAspAsnSerSerProProAlaValProGlnSerTyGlnValAlaHis 202
 DB 279 CCGCGTGTCCATTCGAGAGAGACTCCTGCTCTGTAATTCGAATTCACCAAGTCTCAT 338
 QY 202 sleuHisAlaProThnGlySerGlyLysSerThrLysValPro-----Al 217
 DB 339 CATTAAGCGCAGACAGGCTCAGGAAGACCAACCATGATCCCGCATATCTCTTGAGGA 398

QY 217 aAlaTyAlaAlaGlnGlyTyTyLysValLeuValLeuAsnProSerValAlaAlaThrMe 237
 DB 399 GGGTTATACAAACAAGGTATGAGATGCTGACCAACCCGAGAGTGGCTGACAT 458
 QY 237 tGlyPheGlyAlaTyMetSerLysAlaHisGly-----IleAspPr 251
 DB 459 GAGTGTGCGCCGCCAGTGGCCGGAGATGGTGTGAAGCTTGGGAATGAGGTTGGCTA 518
 QY 251 oAsnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleTh 267
 DB 519 CAGCATCCGCTTTGAGAGCTGACATCAAGAGAACTCTC-----CTCCG 563
 QY 267 rTySerThrTyGly-----LysPheLeuAlaAspGlyGlyCySerGlyG1 283
 DB 564 CTACATGACAGATGGAGATGCTTCCGGAGATCTCTCTGCTGAGCTGACCTGGC----- 618
 QY 283 yAlaTyAspIleIleIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuG1 303
 DB 619 -AGTTACAGCGGTGATGTGTGATGAGCAGACACAAAGACCTTACACAGACATTCCT 677
 QY 303 yIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaTh 323
 DB 678 CTTTGATTCATCAAGATGTTGCTCGCTC----- 708
 QY 323 rAlaThrProProGlySerValThrValProHisProAsnIleGluGluValAlaLeuSe 343
 DB 709 -----CGACCTGAGCTCAAGTCTGTGGCTTC 737
 QY 343 rThr-ThrGlyGluIleProPheTy----- 351
 DB 738 AGCAATGACACATCCCGCTTTTCCACTTCTTGATGAGCCCTGTGTTGATTC 797
 QY 352 --GlyLysAlaIleProLeuGlnAlaIleLysGlyGlyArgHisIleuIlePheCysHis 371
 DB 798 CCGAGACGAGGTTCTCTGTGAC-----ACTTCTACACCA 833
 QY 371 erLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyValaAsnAlaVala 391
 DB 834 AGGCTCCAGAGGCTGACTTGAAGCTTGTGTATATCTGTG----- 877
 QY 391 laTyTyArgGlyLeuAspValSerValIleProThrSerGlyAspVal 407
 DB 878 -----TTGCAGATCCATGTGAC--CGACCCCTGGGGATATC 912
 RESULT 9
 AK031534
 LOCUS AK031534 2972 bp mRNA linear HTC 18-SEP-2003
 DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone:603048F20 product:hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein, full insert sequence.
 ACCESSION AK031534
 VERSION AK031534.1 GI:26327388
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 REFERENCE 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2
 10349636
 TITLE Carninci, P. and Hayashizaki, Y.
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL MEDLINE
 PUBMED 20499374
 11042159
 3

AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
11076861

TITLE
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
REFERENCE
AUTHORS
6
(bases 1 to 2972)
Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Adachi, J., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 2972
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:6030448P20"
/db_xref="MG1:2394772"
/db_xref="taxon:10090"
/clone="6030448P20"
/sex="male"
/issue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
17. 2641
/note="unamed protein product; hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOP|52540, evidence: SCOP)
putative"
/codon_start=1
/protein_id="BAC27438.1"
/db_xref="GI:26327389"

FEATURES
source

CDS

/translation="MAAGRVGPSPRPPTMAWPKKRLRRRTGASQGRSDSDGSE
FEIQAEADAAARKGPSPRALPSPTSCVSDVEDEMEPAQKKKKSGSGFQSMGL
SYVFKGIMKKGVVPPPIORRTIPVLIDKDVAMRTSGKTAFLPMEFLKAR
SAQNGARLILSPREKALQMTKEKLEKRTGLKTLILIGDGMEOFLALNPDI
IATPFGLVHVAQEMNKLOSVEVYVDEADRLPEMGPABLOQITIGRPGHDTVP
SATLPKLVEPAQKGLTEPVLITRDVDSKLEQIKTSFPLVRETPKAVLYLQNV
RPQOQTVPAATKHAAYLTELMDQVSCAHITYSALDQTKNTLAKFTNKSTLI
VTDLAAGLIPLDININYSFPAKGLFLHRVVARAGSGTASLVADDEVYLL
DLHLFLRSYTLARCEPSVADAVGRDGVLRVQSVDEDSLOTAMGASILDG
LHRVANAQOQVRSRPAPESTIKRAKELDLAELGHPLESCFEEGELQRLYDS
IKATRTTTTETIYASSKSDPSQMRKAKRORDRAVAFQQRBERORBPAPAPRE
LPQEESEWETVGEVTEVQRPSPQSGAKRRMETRQDQETVYRPFDP
SERGLSVSGAGAFQOQVAGVLDLMDQENMSRQQLKMDRKKRFPVQSQEDK
KKTIESGRTFSSSKYKRDLYQMKOKKIDIRDBEESGPSRQRPGRGKGRSGQ
TSQRPASSVPAGRMRSKLTKEQLKQRPQAKOKRQFLQRLKQLSANRRRAQELRQ
GARGRGAPSRKGKRRKM"

2954. 2959
/note="putative"
2972
/note="putative"

polYA_signal
polYA_site

ORIGIN

Alignment Scores:
Pred. No.: 3.47 Length: 2972
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservative: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match: 3.57% Indels: 201
DB: 11 Gaps: 29

US-09-930-591-2 (1-686) x AK031534 (1-2972)

QY 100 CysglYserSerAapLyuTyrLeu-----ValThrArgHlaAlaSpVal 114
Db 90 TGGCAAAACGGCGAAGCTGGGCTTCCAGAGCCCGACAGGAGCATCGGATGACGGCGAGT 149
QY 115 IleProValArgArgAG-----GlyAspGlyArgIySerLeuLeuSerProArg 131
Db 150 TCGAGATCCAGCGGAGGATGACCCCGGGAGAGAGCTGGGCTCGAGAGCTTGC 209
QY 132 Pro-IleSerTyrLeuYs-----GlySerSerGlyProLeuLe 145
Db 210 CCTCATTTCCATCCTCAGAGGCTGATGATGAGCCCGACATCGGAGAGATGCTGC 269
QY 145 uCyProAlaGlyHisAlaValGlyIlePheArgAlaVal----- 159
Db 270 GAGCCCAAGAACAAAGAAAGAGTCTGAGGCTTCCAGTCCATGAGGCTGAGTTACC 329
QY 160 -CysThr-ArgGlyValAlaValAlaValAerPheIleProValIgluSerLeuGluThrT 179
Db 330 CTGTCTTCAAGGAGATGATAAAGGCTTCAAGGTGCCAGCCCATCCAGAGAGAA 389
QY 179 hrMetArgSerProValPheSerAerPansSerSerProProAlaValProGlnSerTyrG 199
Db 390 CCATC-----CCGTAATCTTGATGCAAGAGATGTGTGGCCATG----- 430
QY 199 InvAlaHisLeuHisAlaProThrGlySerGlyYysSerThr----- 213
Db 431 -----GCCCGACAGGCAAGTGGCAAGAGCGGCTGCTCTCTCCCGA 473
QY 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAerP 231
Db 474 TGTTTGAGCGGCTAAGGACGACGAGTGCACAGCGGGGCTCGAGCCCTCATCTCTCAC 533
QY 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisG 248
Db 534 CCACCCGAGAGTGGCCCTGACAGCATGAAGTTCATTAAGAGCTAGGCAAGTT-CACC 592
QY 248 Ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrGlys 264
Db 593 GGCCTCAAGACTGCCTGATCTCGGTGGAGACAAAATGAAGACCAAGTTGACGCCCTG 652
QY 264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAerGlyGlyCysSerGlyGlyA 284

Db	653	CACGAGAACCTGACATATATCATTTGGCCACCCTGGGGGCTGTGGATGTG----	GGCT	706
Qy	284	LaTYrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI	---	304
Db	707	GTGGAGATGAACCTTGAAGCTGCAGAGTCT-----	GGAGTATGTGGTG	748
Qy	304	Le-----	---	304
Db	749	TTTCATGAAGACAGACGCTCTTTGAATGGCTTTGCTGAGACGACTACAGAGATCAT	A	808
Qy	305	-----GlyThrValIleAsp-----GlnAlaGluThrAlaG	---	315
Db	809	GGCGGCTTCTGGGGGCGACCAACGGTGTGTCTTCACTACACTGACCTGCCAAGCTGCTG	---	868
Qy	315	LYAlaArgIleuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV	---	332
Db	869	GTGGAATTTTGACACAGGACGAGCTCAGAGAGCCGTGCTCATCCGCTGAGAGTAACTCC	---	928
Qy	332	aIProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIlePro-----	---	349
Db	929	AAGTCATATGACAGCTCAAGACCTC-----	CTTCTCTCTGTG	967
Qy	350	-----PheTyrglyLysAlaIleProLeuGluAlaIleLys-----GlyG	---	363
Db	968	CCCGAAGACACCAAGCTGCTGTGCTCTTCACTGCTGACAGATGTGTTGGCCCCAG	---	1027
Qy	363	ly-ArgHisLeuIlePheCysHisSerIleLysIleCysAspGluLeuAlaIleLysLeu	---	382
Db	1028	GACCCAGACTGTGGTGTGTTGAGCCACAAAGACCATGCGAGTACCTCAGAGACTGTG	---	1087
Qy	383	ValAlaLeuGlyValAsnAlaValAlaIleTyrglyLeuAspValSerValIlePro	---	402
Db	1088	ATGGGCCAGGGGTGATGTTGGCCCAACTATATGTCCTTGACACAGAG-----	---	1138
Qy	403	ThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThrGlyAsp	---	422
Db	1139	-----GCCCGCAAGATCACTGGCCCAATTGACACACAC	---	1174
Qy	423	PheAspSerValIleAspCysAsnThr-----CysValThrGlnThrValAspPheSerLeu	---	441
Db	1175	-----AAATGTTCCACCTCATCGTACCTGACCTGGCCGCCGGGGCTG	---	1219
Qy	442	AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla-----	---	456
Db	1220	GACATCCCACTGTCGACACAACGTCACTACACTTCCTCCGCCAAGGAGACTCTTC	---	1279
Qy	457	-----ValSerArgThrGlnArgArgGlyArgGlnArgGlyLysProGlyLe	---	473
Db	1280	CTGCACCGAGTGGCGCTGTGGCCCGAGACGGCCCAAGTGGACAGCC-----	---	1327
Qy	474	TYrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerValLeuCy	---	493
Db	1328	TATCTCTTGTTGGCCCAAGACGAGGTCCCTTACC-----	---	1361
Qy	493	seGluCysIlyrAspAlaGlyCysAlaIleProIlyGluLeuThrProAlaGluThrValAr	---	513
Db	1362	-----TGCCTTGACCTACACCTGTTCCTGGGCGCGCTGTGACCTTGACCCGCTCTTGTAAG	---	1417
Qy	513	G---LeuAlaIleAlaTyMeLeuAsnThrProIlyLeuProValCysGlnAspHisLeuGluPh	---	532
Db	1418	GAGCTTTCAGTGGCAGATCGGTGGCAGAG-----ACGAGATGG	---	1455
Qy	532	eTTPGluGly-----	---	535
Db	1457	CTGGGTGCGGTGCCCGAGATGTAGTATGAGACAGACAGACCTGACAGCTCCATG	---	1516
Qy	536	-----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGluTh	---	551
Db	1517	GGGGCATCCCTGGATCTTCAGGGCTGCAC-----	---	1544
Qy	551	rlyGlnSerGlyGluAsnLeuProTyrglyLeuValAlaTyrglnAlaThrValCysAlaAr	---	571

Db	1548	-----	GCCTGGCCACAACCGCTCAGACGAGTATGTGCGCTCA	1585
Oy	571	gAlaGlnAlaPro	-----	ProPserTyrPsp---GlnMetTyr 583
Db	1586	CGGCGACGGCCCTGGCCTGATGCATCATGAGAGCCAAAGAGCTGACCTTGACAGACTG	1645	
Oy	583	pLycCys 585		
Db	1646	GGCTTGC 1652		
RESUT.T 10				
AK031679				
LOCUS				
DEFINITION	AK031679	2984 bp	mRNA	linear
ACCESSION	AK031679.1	GI:26327530		
VERSION	AK031679.1	GI:26327530		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Yamawake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichipillary sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	20530913			
PUBMED	11076861			
REFERENCE	4			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature 409, 685-690 (2001)			
MEDLINE	11076861			
PUBMED	11076861			
REFERENCE	5			
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs			
JOURNAL	Nature 420, 563-573 (2002)			
MEDLINE	120530913			
PUBMED	120530913			
REFERENCE	6 (bases 1 to 2984)			
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohse, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shikata, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,			

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Submitted (16-JUL-2001) Yoshihide Hayashizaki. The Institute of
Direct Submision

Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehito-cho, Tsukuba-shi, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers

FEATURES

SOURCE

CDS

```
1..2984
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6J"
  /db_xref="FANTOM_DB:6030486C03"
  /db_xref="WGI:2394945"
  /db_xref="taxon:10090"
  /clone="6030486C03"
  /sex="male"
  /tissue_type="testis"
  /clone_id="RIKEN full-length enriched mouse cDNA library"
  /dev_stage="13 days embryo"
  /note="unnamed protein product; hypothetical P-loop
  containing nucleotide triphosphate hydrolyases structure
  containing protein (SCOP|52540, evidence: SCOP)
  putative"
  /codon_start=1
  /protein_id="BAC27509.1"
  /db_xref="GI:26327531"
  /translation="MAAGRRVGPSPPTMAPPKKRLRKRRTGASQGRSDSDSE
  FEIQEDDARARLKPGLALPSPFTSECVSEVDTEBWAQNNKKKSGSGFSGWGL
  SYPVFKIMKKGVPTPIORRTIPIVLDGKVAMARSGKACFLPMPERLKR
  SAQAGARALILSPRELALQTMPTKEIGFTGLKTLALIGDIMEBQFALHNPDI
  IATRGRLVHAVMNLKQSEYVVEDEARLPEMGAELQJLIGRLPGHQTVF
  SATLPKLVEPARAGLTPVILRLDNDKINLQKTSPLVRENTKAVLLYLQNV
  RPQDTVPVATKHAETLRLMLGGYSCNHTISALDQTKRKNLAKFTNKSSTLI
  VTDLAAGLDLPILDNVIINYSFPAKGKLPVLRVGVVAPAGSGTAYSLVADVPYLL
  DLHLFLRSVTLARPCEPSVADAVADGVGVAPQSVVDEDESLQVAMGASLDLG
  LHRVANNACQOYVRSRPAESPESIKRAKELDLALGLHPLFSFCFEGEGLRLVDS
  IKVYRTTTFEINASSKDPSSQMKRRAKORRKAIVAFPOORROREGPADPAPRE
  LPOEEEMVETVGVCTVEVVGQRRPRPGSQAKRRMTRORDDQEFVYPRKPD
  SERGLSVSGAGAEQVAVGANVLDLMDGEMNSRGQQLMDKKRFGQSQDEBK
  KKITESGRFTISSYKRDLYQWKQKQKIDRDESEBSPNQRPGRPGGRKRSQ
  TSGPRASSVPAGRMRSSELTKEQILKQPGQSEALPAARGPAAFTSTPPSPGAAP
  GRLLPGCSLQGGQDEEKDVRSQTQPMGFLVAVGVDRDYYVHLLDEPSTGAL"
  2963..2968
  /note="putative"
  2984
  /note="putative"
  2984
  /note="putative"
```

ORIGIN

Alignment Scores:

Pred. No.: 3.5 Length: 2984
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservative: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match: 3.57% Indels: 201
DB: 11 Gaps: 29

US-09-930-591-2 (1-686) x AK031679 (1-2984)

QY 100 CysGlySerSerAspLeuTyrIleu-----ValThrArgHisAlaAspVal 114
|||||.....|

```
Db 98 TCGCAAAAGCCGAACTGGGGCTTCCCAAGCCGCGACAGCGACTCGGATGACGGCGACT 157
Qy 115 lIleProValArgArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
Db 158 TCGAGATCCAGCGCGAGGAGTACCGCCCGGCGAGAGAGCTGGGCTTCGAGAGCTTGC 217
Qy 132 Pro-IleSerTyrIleuLeu-----GlySerSerGlyGlyProLeuLe 145
Db 218 CTTCAATTCCTACCTCAGAGTGGCGTATCATGATGTGAGCCCGACATCGGAGAGATGTGTC 277
Qy 145 uCyProAlaGlyHisAlaValGlyIlePheArgAlaVal----- 159
Db 278 GAGCCGAAACAAGAAAAGAGAAGAGTGGAGCTTCAGCTCATGAGCCCTGATTACC 337
Qy 160 -CysThr-ArgGlyValAlaValAlaValAlaPheIleProValGlySerLeuThrT 179
Db 338 CTGTTTCAGGAGGATATGAAAAAGGCTTACAGAGTCCGACGCCATCCAGAGAA 397
Qy 179 hMetArgSerProValPheSerAspAnSerSerProProAlaValProGlnSerTyrG 199
Db 398 CCATC-----CCGTCATCTTGATGCAAGAGATGTGGCCATG----- 438
Qy 199 lNValAlaHisLeuHisAlaProThrGlySerGlyIleSerThr----- 213
Db 439 -----GCCCGGACAGGAGTGGCAAGAGCGGCTGCTTCTCCCGA 481
Qy 214 -----LysValProAlaAlaValAlaValAlaGlnGlyTyrValLeuValLeuAnp 231
Db 482 TGTTTGAGCGGCTGAAAGCAGCGATGCAAGAGCGGGGCTTCAGCCCTCATCTTCAC 541
Qy 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerIleHis 247
Db 542 CCACCGGAGAGTGGCCCTCGACAGCATGAAGTTCACTAAAGACTGAGCAAGTT-CACC 600
Qy 248 -----GlyIleAspProHisIleArgThrGlyValArgThrIleThrThrGlyS 264
Db 601 GGCCTCAAGACAGGCTTGATCTGGGTGGAGACAAATGAGACAGATTGACGCCCTG 660
Qy 264 eProlIethrTyrSerThrTyrGlyIlePheLeuAlaAspGlyGlyCysSerGlyGly 284
Db 661 CACGAGAACCTGACATATATATTCGACCCCTGGGCTCTGGTGCATGT-----GGCT 714
Qy 284 lArYAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSerIleLeuGlyI 304
Db 715 GTGAGAGTGAACCTGAAGCTGACAGAGTGT-----GAGATGATGGTG 756
Qy 304 le----- 304
Db 757 TTCATGAGACAGACAGGCTCTTGAATGGGCTTGTCTGAGCAGCTACAGAGATGATA 816
Qy 305 -----GlyThrValLeuAnp-----GlnAlaGlyThrAlaG 315
Db 817 GGGCGCTTCTGGGGGCGACACAGAGTGTCTGTTCTCAAGCTACATGCCCCAAGCTGTG 876
Qy 315 lYAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
Db 877 GTGGAATTTGACCGGAGGAGGCTCAGAGGCGGTGTCATCCGCTGAGAGTGAAGTACC 936
Qy 332 aIProHisProHisIleGlyGlyValAlaLeuSerThrThrGlyGlyIlePro----- 349
Db 937 AAGCTCAATGAGCAGCTCAAGACCTC-----CTTCTCTCTTGTG 975
Qy 350 -----PheTyrGlyIleValAlaIleProLeuGlyAlaIleLeu-----GlyG 363
Db 976 CGCGAAGACCAAGGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1035
Qy 363 lY-ArgHisLeuIlePheCysHisSerIleLeuLeuCysAspGlyLeuAlaValLeu 382
Db 1036 GACCAAGCTGTGTGTTTGAAGCACAAGACATGCGAGTGTCTCACAGAGTGTCTG 1095
Qy 383 ValAlaLeuGlyValAlaAlaValAlaValTyrTyrArgGlyLeuAspValSerValIlePro 402
Db 1096 ATGGGCCAGGCTGTGAGTGTGGCCACATCTATATGTCCTTGGACCAAGG----- 1146
```

QY 403 ThrseryGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
 Db 1147 -----GCCGCAAGTCACTGGCCAGATTCCACACACAC 1182
 QY 423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
 Db 1183 -----AAATGTTCCACCTCTATCTGTGACTGACCTGGCCGCCGGGGCTTG 1227
 QY 442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
 Db 1228 GACATCCCACTGCGGCAACAGCATACACAGCTTCCCTGCCAAGGCAAGCTCTTC 1287
 QY 457 -----ValSerAgtThrGlnArgArgGlyArgThrGlyArgGlyArgProGlyIle 473
 Db 1288 CTGACCCAGAGTGGCGCGTGGCCCGAGCAGCGCCGAAGTGGCACAGCC----- 1335
 QY 474 TyrArgPheValAlaProGlyGlyu-ArgProSerGlyMetPheAspSerValLeuCy 493
 Db 1336 TATCTCTTGGTGGCCCGACAGAGTCCCTAC----- 1369
 QY 493 sglucyTyraAspAlaGlyCysAlaTyrGlyLeuThrProAlaGluThrThrValAr 513
 Db 1370 -----TGCTTGACCTACACCTGCTGGCGCGCTGTGACCTGGCCCTTGTGTAG 1425
 QY 513 g---LeuArgAlaTyrMetLeuThrProGlyLeuProValCysGlnAspHisLeuGlnPh 532
 Db 1426 GAGCCTTCAGTGGCAGATGGCGTTGGCGAGG-----ACCGAGTGC 1464
 QY 532 eTyrGlnGly----- 535
 Db 1465 CTGGTCCGCGTGGCCCGACAGTGTAGTGTATGATGAGACAGACGCTGCAGATGCCATG 1524
 QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
 Db 1525 GGGGCGATCCCTGAGCTTCAGGGCGCTGCAC----- 1555
 QY 551 rlyGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
 Db 1556 -----GCGTGGCCAAACAGCTCAGCAGCACTATGTGGCTCA 1593
 QY 571 gAlaGlnAlaPro-----ProProSerTyrPasp---GlnMetTr 583
 Db 1594 CGGCGACGGCGCTCGCTGATGTCATCAAGAGCCAGAGAGCTGAGCTGGCAGAGCTG 1653
 QY 583 pLyScy 585
 Db 1654 GGCTTGC 1660

RESULT 11
 AK028274 2986 bp mRNA linear HTC 18-SEP-2003
 LOCUS Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
 DEFINITION enriched library, clone:3222401E18 product:hypothetical P-loop
 containing nucleotide triphosphate hydrolases structure containing
 protein, full insert sequence.
 AK028274
 ACCESSION AK028274.1 GI:26080791
 VERSION AK028274
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to

JOURNAL
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE
 AUTHORS 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Komuro, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichipillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL 20530913
 MEDLINE 11076861
 PUBMED 11076861
 REFERENCE
 AUTHORS 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE
 AUTHORS 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 120530913
 PUBMED 120530913
 REFERENCE
 AUTHORS 6
 (bases 1 to 2986)
 Raichi, T., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komuro, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN) Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1..2986
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM:DB:3222401E18"
 /db_xref="MG1:2350122"
 /db_xref="taxon:10090"
 /clone="3222401E18"
 /tissue_type="head"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="14, 17 days embryo"
 21..2751
 /note="hypothetical P-loop containing nucleotide
 triphosphate hydrolases structure containing protein
 (SCOP|52540, evidence: SCOP)

misc_feature

QY	442	AspProthrPheThrIleGluThrIleThrIleuProGlnAspAla-----	456
Db	1210	GACATCCACCTGCTGGAGCAACGCTTCAACTACAGCTTCCCTGCCAAGGAGAGCTTTC	1263
QY	457	-----ValSerArgThrGlnAspArgIleArgThrIleArgGlyIleProGlyIle	473
Db	1270	CTGCACCGAGTGGGCGCGTGTGGCCCGCCAGACAGGCGCAAGTGGGACAGCC-----	1317
QY	474	TyrArgPheValAlaIleProGlyIleu-ArgProSerArgIlePheAspSerValLeuCy	493
Db	1318	TATCTTGTGGTGGCCCGCCAGACGAGGTCCCTTAC-----	1351
QY	493	SeLUCysIleThrAspAlaGlyIleAlaIlePTrIleIleuThrProAlaGluThrIleValAr	513
Db	1352	-----TGCTTGACCTTACACCTGTCTTCTGGCGCGCTCTGTACACCTGGCCCGCTTGTAG	1407
QY	513	G-----LeuArgAlaIleTyrMetAsnThrProGlyIleuProValCysGlnAspHisIleuGluPn	532
Db	1408	GAGCCTTCAGTGGGACGATCCGTTGGCAGGG-----ACGAGTGG	1446
QY	532	eTrpGluGly-----	535
Db	1447	CTGGGTGCGTGGCCCGCCAGAGTGTAGTGTATGATGATGAGACAGACGCTGCAGCTGCCATG	1506
QY	536	-----ValPheThrGlyIleuThrHisIleAspAlaHisPheLeuSerGlnTh	551
Db	1507	GGGCGATCCCTGGATCTTTCAGGCGCTGCACC-----	1537
QY	551	IleGlnIleSerGlyGluMetAsnIleuProTyrIleValAlaIleThrIleValCysAlaAr	571
Db	1538	-----GCGTGGCCAAACGCTCAGCAGCAGCATATGTGGCTCA	1575
QY	571	GAlaGlnAlaPro-----ProProSerTyrAsp---GlnMetTr	583
Db	1576	CGGCGACGAGCCCTCGCTGAGTCCATCAAGAGAGCCAGAGAGCTGGACCTGGCAGAGCTG	1633
QY	583	PLVSeCys	585
Db	1636	GGCTTGC	1642
RESULT 13			
LOCUS	AY404177	1697 bp	DNA linear
DEFINITION	Mus musculus DDX34 gene, VIRTUAL TRANSCRIPT, partial sequence,		GSS 15-DEC-2003
ACCESSION	AY404177		
VERSION	AY404177.1		
KEYWORDS	GSS.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 1697)		
	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,		
	Todd,M.A., Tenenbaum,D.M., Clavello,D.R., Lu,F., Murphy,B.,		
	Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,		
	Adams,M.D. and Cargill,M.		
TITLE	Interferring nonneutral evolution from human-chimp-mouse orthologous		
	gene tritos		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 1697)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,		
	Todd,M.A., Tenenbaum,D.M., Clavello,D.R., Lu,F., Murphy,B.,		
	Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,		
	Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,		
	Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering		
FEATURES	Location/Qualifiers		

```

source
1. .1697
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1. .>1697
/gene="DDX34"
/locus_tag="HCM1803"

ORIGIN

Alignment Scores:
Pred. No.: 1.44 Length: 1697
Score: 128.50 Matches: 134
Percent Similarity: 35.56% Conservative: 68
Best Local Similarity: 23.59% Mismatches: 221
Query Match: 3.55% Indels: 148
DB: 29 Gaps: 30

US-09-930-591-2 (1-686) x AY040177 (1-1697)

QY 52 ValCystrPrThValTYrHisglValaglYthRarGhrllleAlaserProlysglyPro 71
Db 116 GTCTCTTGGAGAGTATTCCTCCGCG--ATGAGATTAATATC-----GTC 166
QY 72 ValIleGln-----MetYrThRAsnValaAspGlnAspLeuVal 84
Db 161 GTGATTCAGAGAGGTCTCAGAAgTTCGAGCTTCTTGAACGCTCGAGAGATTCCAGC 220
QY 85 GLYrPrAlaAProGlnlYAlaAspSerleuThrProCythrCysglySerSerAsp 104
Db 221 ACCTCAAGACNCTCCAGAAAGAAAGAAAGACNCTCGCATGCCCAACATGCGATGCTG 280
QY 105 Leu-----Tyrleu--ValThrArghIsalaAspValIleProValArGaArgArg 120
Db 281 CTTTAgCTGACNACTCTCACTTCACTTCAAGCCACGCT----- 316
QY 121 GLYAspGlyArgglySerleuSerProArgrProIleSerTYrleuYrsglySerSer 140
Db 317 --ACGCATCAACCTTCATCTTCATCCACCCAGACACTCGGGGTGTCATGAGGCCGCGCA 373
QY 141 GLYGLY-----ProleuSerCyAspProAlaglyHisAlaValIleGlyIlePheArgAla 158
Db 374 GGGGGGTGCCCCCGAGAGAGTGTCCAGATTC--GCCGTGCArTTCCTTCACTCACTAG 430
QY 159 ValCystrArGlyValAlaLysAlaValaAspPheIleProValGlnSerleu--Gln 177
Db 431 ACTTCACAGAGA-----AGCAGGCTTTTGAAGGCTGCCAAGCTCAGAGGTAGCGAG 484
QY 178 Thr-ThrMetArGserProValPheSerAspAsnSerSerProProAlaValProGln-S 197
Db 485 CGGACACTTCCATCGCCGCCAGTAGGGAACCCGAT-----CTTCAAGACTCTCAAG 535
QY 197 erTYrGlnValAlaHisleuHisAlaProThrGlnlySerGlyLysSerThryLeuValPro 217
Db 536 AACACCAAGGtGGTGGTGGTGGCCGGGAGACAGAGGCTGTGGCAATCTACTACAGGTGCC 555
QY 217 laaIatYrZlaaIagInglyTYrLys---ValleuValleuAsnPro----- 231
Db 596 AGTACTGTCGGGCGCGCGGCTTCACTACAGTGGCATGATGATCAAGCCCGAGATGCGCT 655
QY 232 ----SerValAlaAlaThrMetGlyPheGlyAlaTYrMetSerLYaHisGlyIleA 250
Db 656 GCATCTCAGCTGACCAAGACCGCTTGGCTTCGAGAGCTCAGTCAAGTATGAGCTCCAGGTTG 715
QY 250 sPrAsnIleArghThrGlnlyValaGhrTrllleThrThnlySerProIleThrTYrSerT 270
Db 716 GCTTACAGATCCGCGCTT--GAAAGCACTCGGTGGCGAGCCACCAAGATGTGTCTCTGA 772
QY 270 hrTYrGlyLYAspPheleu-----AlaAspGlyGlyCySerGlyGlyAlaTYrAspIleT 288
Db 773 CGGTGGGGGTGCTCTTACGCCAGATCCAGCGTGAAGCCAGCGTGGCCCACTACAGAGTCC 832
QY 288 lelleCyAspGlyCyshISerThrAspAlaThrSerlleuGlyIleGlyThrVal 308

```

```

Db      833 TGATCTGACGAGCTCCACGAGCAGCACTGCACATGACTTCTGCTGGT----- 885
QY      308 euaBpGlnalagluthrAlaGlyAlaArgLeuThrValLeuAlaThrAlaThrProProG 328
Db      886 -----GTCCTGACGAGCGCTGCTGCC----- 906
QY      328 lyeServalThrValProHisProAsnIleGluGluValAlaLeuSerThrThyGlyGluI 348
Db      907 -----CAGCGGCTTACCTCAAGTCATGTCCTCATCTCGGACCATCAACA 952
QY      348 lPProPheTyrGlyAlaAlaIleProLeuGluAlaIleGlyGlyAlaArgHisLeuIleP 368
Db      953 TCTCTCTCTCTCCAGC-----TACT 973
QY      368 heCyHisSerLysLysLysCyAspGluLeuAlaAlaLysLeuValAlaLeuGlyAla 388
Db      974 TCAGCCATGCT-----CCGGTGGTGAGTGAGCTCGGAGGCTTCCCATCAAGGCT 1027
QY      388 snAlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValV 408
Db      1028 TCAGATGGCT-----CCAGCTGAGCTCCGGAAGT 1057
QY      408 alValAlaAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp---PheAspSerValI 427
Db      1058 GCATCTCTCCACCAACATGGCCGAGCATCTGATCAGATGAGCATCCGCTTCTGCTGG 1117
QY      427 leAspCyAsnThrCyValThrGlnThrValAspPheSerLeuAspProThrThrTri 447
Db      1118 TCGATTCTGG-----AAGGTGAAGAGATGAGTACGACCCACAGCCAAAC 1165
QY      447 leGluThrIleThrLeuProGlnAsp-----AlaValSerArgThrGlnArg 463
Db      1166 TGCACCGGCTG-----CAGGAATCTGATCAGCAGCCAGCTGTAGAGAGGCCA 1216
QY      463 rGgLYArgThrGlyArgGlyLysProGlyIle---TyrArgPheValAlaProGlyGluA 482
Db      1217 AGGGCCGAGCGGCGCAGCGGGTCCGGCTGCTACCGTCTCATGCGCAGTGTGAC- 1275
QY      482 rGProSerGlyMetPheAspSerSerValLeuCyGluCySerThrAspAlaGlyCySAlaT 502
Db      1276 -----TATATGCTT-TGGCC- 1292
QY      502 rPTyGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThrProG 522
Db      1293 -----GTACCCAGTCCCG 1305
QY      522 lYLeuProValCySglnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeuThrH 542
Db      1306 GAATTCGG-----CAGATGAA-GAGTATGAGTGGGGGAGCCACGCG 1346
QY      542 lGllAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLysLeu 562
Db      1347 ACCCTTCCTTCATGACGCCCTCCACTGCGAGGTGGAACCGCTATCTCT-ATCT 1405
QY      562 alAlaTyrGlnAlaThrValCySAlaArgAlaGlnAlaProPro-----ProSerTPA 580
Db      1406 CCAAGAGCAGGGGCGCTGAGACGCTCGAGGCGCTCACCCCATCGGGTCCCTGCTGGC 1465
QY      580 spGlnMetTrpLysCySLeu 586
Db      1466 CC-----AGCTGCTTG 1476

```

```

RESULT 14
BU054966      691 bp      mRNA      linear      EST 26-AUG-2002
LOCUS      UI-M-FD0-bzi-n-19-0-UI-r1 NIH_BMAP_FDO Mus musculus cDNA clone
DEFINITION      IMAGE:6404178 5', mRNA sequence.
ACCESSION      BU054966
VERSION      BU054966.1 GI:22495043
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
1 (bases 1 to 691)
AUTHORS      NIH-MGC http://img.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgabs-r@mail.nih.gov
      Tissue Procurement: Dr. James Lin, University of Iowa
      cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
      DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
      Clone Distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/BLN at:
      http://image.llnl.gov
      This clone was contributed by the Brain Molecular Anatomy Project
      (BMAP)
FEATURES
source
Location/Qualifiers
1..691
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6404178"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH BMAP_FDO"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGACC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene discovery in the developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      0.369      Length:      691
Score:      127.50      Matches:      45
Percent Similarity:      44.19%      Conservative:      31
Best Local Similarity:      26.16%      Mismatches:      69
Query Match:      3.52%      Indels:      27
DB:      13      Gaps:      7

```

US-09-930-591-2 (1-686) x BU054966 (1-691)

```

QY      174 GluSerLeuGluThrThrMetArgSer---ProValPheSerAspAsnSerSerProPro 192
Db      145 GAATCTATTCACAGCGCTGCGCGCGCTGCGCTGCTTCCCTTCCAGAGGACTTCTG 204
QY      193 AlAlaValProGlnSerTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSer 212
Db      205 GCCGCCATTGCCAACATCAGTCTCTCATTCGAAGCGAGCATGGCTCGGAGAAGCC 264
QY      213 ThrLysValPro-----AlaAlaTyrAlaAlaGlnGlyTyrLysValLeu 227
Db      265 ACAAGATCCACAGTACCTTTGAGAGAGGTTACACAAAGAGGGATGAAATATGCT 324
QY      228 ValLeuAsnProSerValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis 247
Db      325 TGCACCCAGCCCGGAGAGTGGCGGCTTATGAGTGGCAGGCCAGATGGCCGGAGATG 384
QY      248 Gly-----lLeAspProAsnIleArg-----ThrGlyVal 257

```


Db 385 GGTGTGAAGCTTGGAGAGAGTGGCTACAGATCCGGTTGAGACTGCACCTCAGAG 444
 Qy 258 ArgThrIleThrThrGlySerProIleThrTySerThrTyGly-----Lys 273
 Db 445 CGAAGTGT-----CTCCGCTACAGACAGATGAGATGCTACTCCGAGAG 489
 Qy 274 PheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyraAspIleIleCysAspGlyCys 293
 Db 490 TTCCTCTCTGAGCTGACCTTGCA-----AGTTACAGTGTGATGATGATGATGAGCT 543
 Qy 294 HisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaIuThr 313
 Db 544 CACGAGCGGACCTTGACACAGACATCTCTTTGATTGATTCAAAGACGTGCTAGATTC 603
 Qy 314 AlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
 Db 604 CGACCTGAGCTCAAGTCTGTGTGCTTCAGCCACA 639
 RESULT 15
 CD240900 738 bp mRNA linear EST 22-MAY-2003
 LOCUS AGENCOURT 14124904 NIH MGC 177 Mus musculus cDNA clone
 DEFINITION IMAGE:30380772 5', mRNA sequence.
 ACCESSION CD240900
 VERSION CD240900.1 GI:31001364
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE NIH-MGC
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsabde-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Michael Brownstein Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.liml.gov
 Plate: NDCM173 row: f column: 13
 High quality sequence stop: 482.
 Location/Qualifiers
 1..738
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30380772"
 /lab_host="DH10B (TI-phage-resistant)"
 /clone_lib="NIH_MGC_17"
 /note="Organ: liver; Vector: pDNR-LIB; Site 1: SfiI
 (ggccatattggc); Site 2: SfiI (ggccgctcggc); cDNA made
 by oligo-dt priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGACAGTGGCATTCAGGCGGG-3' and
 5'-ATCTAGAGCGGAGCGGCGCCACATG-dT(30)NN-3'. Full-length
 enriched library was constructed using the Clontech
 Creator SMART kit and size-selected to contain the 0.5 kb
 size fraction. Library created in the laboratory of M.
 Brownstein (NIH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 0.414 Length: 738
 Score: 127.50 Matches: 45
 Percent Similarity: 44.19% Conservative: 31
 Best Local Similarity: 26.16% Mismatches: 69
 Query Match: 3.52% Indels: 27
 DB: 14 Gaps: 7

US-09-930-591-2 (1-686) x CD240900 (1-738)
 Qy 174 GluSerLeuGluThrThrMetArgSer---ProValPheSerAspAsnSerProPro 192
 Db 62 GAATCTATCCAGGCTGTGCGCGGACGCTGCCCTGTTCCTCCAGAGAGCTTCTG 121
 Qy 193 AlaValProGlnSerTyGlnValAlaHisIleuHisAlaProThrGlySerGlyLysSer 212
 Db 122 GCCGCCATTGGCAACATCAGTGTCTCATCATCAAGGCGAGACTGCTCGGAGAAC 181
 Qy 213 ThrIysValPro-----AlaAlaTyraAlaAlaGlnGlyTyTyValLeu 227
 Db 182 ACACAGATCCACAGATCCTTTGAGAGGGTTACCAAGAGGCGATGAGATGCT 241
 Qy 228 ValLeuAsnProSerValAlaAlaThrMetGlyPheGlyAlaTyMetSerLysAlaHis 247
 Db 242 TGCACCCAGCCCGGAGAGTGGCGCTATGATGTGGACGCGCAGATGCGCGGAGATG 301
 Qy 248 Gly-----IleAspProAsnIleArg-----ThrGlyVal 257
 Db 302 GGTGTGAAGCTTGGGAGAGAGTGGCTACAGATCCGTTTGAGACTGCACCTCAGAG 361
 Qy 258 ArgThrIleThrThrGlySerProIleThrTySerThrTyGly-----Lys 273
 Db 362 CGAAGTGT-----CTCCGCTACATGACAGATGAGATGCTACTCCGAGAG 406
 Qy 274 PheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyraAspIleIleCysAspGlyCys 293
 Db 407 TTCCTCTCTGAGCTGACCTTGCA-----AGTTACAGTGTGATGATGATGATGAGCT 460
 Qy 294 HisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaIuThr 313
 Db 461 CACAGAGGAGACTTGACACAGACATCTCTTGTGATGATCAAGAGCTGCTAGATTC 520
 Qy 314 AlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
 Db 521 CGACCTGAGCTCAAGTCTGTGTGCTTCAGCCACA 556

Search completed: September 17, 2004, 12:08:02
 Job time : 5807 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 06:45:20 ; Search time 7206 Seconds

(without alignments)
4126.189 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618
Sequence: 1 MAPTAYAOOTRGHLCIT.....PAIPDRVLYREFDEWEEC 686

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n model -DEV=xlh  
-Q=/cgm2_1/USFTO.epool/US0930591/runat_13092004_164952_686/app_query.fasta_1.839  
-DB=genbml -QFMT=faetap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -QFMT=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US0930591 @CCN 1.1 4390 @runat_13092004_164952_686 -NCPU=6 -ICPU=3  
-NO_MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEVTIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
1: gb_ba:*  
2: gb_ncg:*  
3: gb_in:*  
4: gb_on:*  
5: gb_ov:*  
6: gb_pat:*  
7: gb_ph:*  
8: gb_pl:*  
9: gb_pr:*  
10: gb_ro:*  
11: gb_sts:*  
12: gb_ey:*  
13: gb_un:*  
14: gb_vl:*  
15: em_ba:*  
16: em_fun:*  
17: em_hum:*  
18: em_in:*  
19: em_mu:*  
20: em_on:*  
21: em_or:*  
22: em_ov:*  
23: em_pat:*  
24: em_ph:*  
25: em_pl:*  
26: em_ro:*  
27: em_sts:*  
28: em_un:*
```

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3618	100.0	2061	6 AX441176	AX441176 Sequence
2	3618	100.0	2061	6 AX467113	AX467113 Sequence
3	3574	98.8	5360	6 I06434	I06434 Sequence 48
4	3574	98.8	5360	6 I09328	I09328 Sequence 8
5	3574	98.8	6785	6 I06440	I06440 Sequence 54
6	3574	98.8	6785	6 I09329	I09329 Sequence 10
7	3574	98.8	7310	6 AR118696	AR118696 Sequence
8	3574	98.8	7310	6 I09331	I09331 Sequence 15
9	3574	98.8	7310	14 HPCPOLYP	M32084 Hepatitis C
10	3574	98.8	9185	6 I08294	I08294 Sequence 1
11	3574	98.8	9185	6 BD091382	BD091382 HCV culti
12	3574	98.8	9379	6 AR166930	AR166930 Sequence
13	3574	98.8	9379	6 AR301300	AR301300 Sequence
14	3574	98.8	9401	6 AR176483	AR176483 Sequence
15	3574	98.8	9401	6 E66593	E66593 Hepatitis C
16	3574	98.8	9401	6 I71894	I71894 Sequence 9
17	3574	98.8	9401	6 I81885	I81885 Sequence 9
18	3574	98.8	9401	6 BD080334	BD080334 Hepatitis
19	3574	98.8	9609	12 HPCPOLYP	M62321 Hepatitis C
20	3574	98.8	9609	12 AF387805	AF387805 Synthetic
21	3574	98.8	9646	12 AF387806	AF387806 Synthetic
22	3574	98.8	9693	12 AF387807	AF387807 Synthetic
23	3571	98.7	6785	6 AR118692	AR118692 Sequence
24	3570	98.7	8316	6 AR118703	AR118703 Sequence
25	3570	98.7	9185	6 AR118722	AR118722 Sequence
26	3570	98.7	9185	6 AR118723	AR118723 Sequence
27	3570	98.7	9618	14 AF271632	AF271632 Hepatitis
28	3568	98.6	6299	6 AX164584	AX164584 Sequence
29	3565	98.5	2058	6 AR404933	AR404933 Sequence
30	3565	98.5	2058	6 AR408362	AR408362 Sequence
31	3565	98.5	2058	6 AX395309	AX395309 Sequence
32	3565	98.5	2058	6 AX454818	AX454818 Sequence
33	3565	98.5	8987	6 AR118728	AR118728 Sequence
34	3565	98.5	9609	12 AF387808	AF387808 Synthetic
35	3555	98.3	5360	6 AR118686	AR118686 Sequence
36	3551	98.1	9646	6 AR110828	AR110828 Sequence
37	3551	98.1	9646	6 BD069982	BD069982 Functiona
38	3551	98.1	12980	6 AF009606	AF009606 Hepatitis
39	3551	98.1	12980	6 AR110831	AR110831 Sequence
40	3551	98.1	12980	6 BD069985	BD069985 Functiona
41	3550	98.1	9395	14 AF511950	AF511950 Hepatitis
42	3550	98.1	9502	6 E08263	E08263 gRNA of Hep
43	3550	98.1	9502	6 E08264	E08264 CDNA of Hep
44	3550	98.1	9502	14 HPCUCJ1	D10749 Hepatitis C
45	3547	98.0	9610	14 HEC27830	AJ27830 Hepatitis

RESULT 1

ALIGNMENTS

AX441176
 LOCUS AX441176 2061 bp DNA linear PAT 28-JUN-2002
 DEFINITION Sequence 16 from Patent WO0213855.
 ACCESSION AX441176
 VERSION AX441176.1 GI:21665758
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.
 REFERENCE
 1 Salberg, M. and Hultgren, C.
 VACCINES CONTAINING RIBAVIRIN AND METHODS OF USE THEREOF
 Patent: WO 0213855-A 16-21-FEB-2002;
 JOURNAL TRIPEP AB (SE)
 FEATURES
 source location/Qualifiers
 1..2061
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Hepatitis C virus NS3/4A coding region"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,41e-203 Length: 2061
 Score: 3618.00 Matches: 686
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-930-591-2 (1-686) x AX441176 (1-2061)
 QY 1 MetAlaProIleThrAlaTYrAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
 Db 1 ATGGGCTTATCAAGGCTATGCCAGACAGAGGGGCTTTGGATGATATAC 60
 QY 21 SerLeuThrGlyArgAspIysAsnGlnValGlnGlyValGlnIleValSerThrAla 40
 Db 61 AGCTTGACCGGCGGCAAAAACAGGTGGAGGTGAGTTCAGATCGTCAACTGCT 120
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTYrHisGlyAla 60
 Db 121 GCCCAGACTTCTTGGCAACTGATTAACGGGGTGTGTGACTGTCTACATGAGACC 180
 QY 61 G1YThrArgThrIleAlaSerProIysGlyProValIleGlnMetTYrThrAsnValAsp 80
 Db 181 GGAACAGAGACCATTTGGCTACCTAAGGTCCTGTTATCCAGATGACCAACAGTGGAC 240
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGTGCCTCCATTAACACCATGACACTTGC 300
 QY 101 GlySerSerAspLeuTYrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 Db 301 GGCTCTCTGGACCTTACCTGATCAAGGACGCGCATGTCATTCTCTGCGCCGACGG 360
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTYrLeuYsgIysSerSer 140
 Db 361 GGTATGSCAAGGGGAGGCTGCTTTCGCCCGGCTATCTTAAAGAGGCTCTCG 420
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheAlaAlaValCys 160
 Db 421 GGAGGCTCTCTGCTGTGCCCGCAGACATCCGTAAGCATATTCAGAGCGCGGTATGC 480
 QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGlnSerLeuGlnThrThMet 180
 Db 481 ACCGTGAGTGGCTTAAGGCGGTGACTTCAATCCCGTAAGAGCTTGAACAACATG 540
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValIleProGlnSerTYrGlnVal 200
 Db 541 AGGTCCCGGTGTCTTCAACAACATCTCCCAACAGACAGTGGCCCAAGCTCAACAGT 600
 QY 201 AlaHisIleuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTYrAla 220

Db 601 GCCCACTGCACTGCCACCGGAGCGGTAAAGACACCAAGTCCCGGCCCAATACGA 660
 QY 221 AlaGlnGlyTYrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 661 GCTCAGGGCTTACAAAGGTGTGGTGTCTCAACCCCTCGTTGCTGCAACAATGGGCTTGGT 720
 QY 241 AlaTYrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 721 GCTTACATCTCCAAAGGCCATGGATTTGATCTTAACATCAGACATGGGGTGAACAATT 780
 QY 261 ThrThrGlySerProIleThrTYrSerThrTYrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 781 ACTACTGGACGCCGATCAAGTATCCACTTCCAGGCAAGTTCTTGGCAGCGGGGTGT 840
 QY 281 SerGlyGlyAlaTYrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
 Db 841 TCAGGGGGTCTTATATACATATATTTGTGACAGTGCATCCACGGAATGCACATCC 900
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 901 ATCTGGCATTTGGCAGTGTCTCTTGCACAAAGACAGACCGCGGGCGAGCTGATGG 960
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGluVal 340
 Db 961 CTGCGCACCGCTACCCCTCGGGCTCCGTCACTGTGCCCATCTTAATCCAGAGGTT 1020
 QY 341 AlaLeuSerThrThrGlyGlyIleProPheTYrGlyLysAlaIleProLeuGlnAlaIle 360
 Db 1021 GCTCTGTCACATCCGAGAGATCCCTTTTATGGCAAGCTATTCCTTGAAGCAATT 1080
 QY 361 LysGlyGlyArgHisIleuIlePheCysHisSerLysLysLysCysAspGlyLeuAlaAla 380
 Db 1081 AAGGGGGGAGACATCTCATCTTGTGCACTGCAATAAGAAAGAGCCACAGCTCGCGCA 1140
 QY 381 LysLeuValAlaLeuGlyValAspAlaValAlaTYrTYrArgGlyLeuAspValSerVal 400
 Db 1141 AAACGTGTGCGTGGCGGCAATGCGGTGCTTACTTACCGGGCTTATGTGTCTGTC 1200
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 1201 ATCCGACACAGTGTACGTTCGTGCGGCAATGACGCCCTCATGACCGGCTTTACC 1260
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrGlnThrValAspPheSer 440
 Db 1261 GCGACCTTCATTCGGTGTAGACTGCCAACACGTGTGTACCCAGACAGTCACTTCAG 1320
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 1321 CTTGACCTTACCTTTCACATTGAGACATACACGCTTCCCAAGATGCTGTCTCCGTACT 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTYrArgPheValAlaProGly 480
 Db 1381 CAACGTGCGGGTGAAGCTGGCAGAGGGAAGCCAGCATCTACAGATTGTGGACCGGGG 1440
 QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGlnCysTYrThrAspAlaGlyCys 500
 Db 1441 GAGCGTCTTCTGCAATGTTGACTGCTGTCTGTCTGCAAGTGTATGACGGGGTGT 1500
 QY 501 AlaTrpTYrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTYrMetAsnThr 520
 Db 1501 GCTTGGTATAGCTTACCGCCCGCAGACACACAGTTAAGCTACGACATACATGAACACC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisIleuGlnPheTrpGlnGlyValPheThrGlyLeu 540
 Db 1561 CCGGAGCTTCCCGTGTGCCAAGACACATCTTGAATTTTGGAGGGCGCTTTAAGGGGTCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGlnAsnLeuProTYr 560
 Db 1621 ACCCACTAAGCCCACTTCTATCCCAACAAGAGTGGGAAACCTTCCCTAT 1680
 QY 561 LeuValAlaTYrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580

Db 1681 CTGGTAGCGTACCAAGCCACCGGTGTGCGTAAAGACTCAAGCCCTCCCGGTGGGAC 1740
 Qy 581 GIMeETrPhyScyLeuileAArgLeuYProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGGAAAGTCTTGATCCGCTCAAGCCCACTCCATGGGCCCAACCTCTCTTA 1800
 Qy 601 TTAArgLeuGlyAlaValGlnAAsngluValThrLeuThrHisProValThrLeuTyrlle 620
 Db 1801 TATAGACTGGGCGGTGTCCAGAAATGAAGTCACTGACGACCCAGTCAACAATATATC 1860
 Qy 621 MeEThrCyMeTsSerAlaAAspleuGluValThrSerThrTrpValIleValGlyGly 640
 Db 1861 ATGACAGTATATGTGCTGACCTGGAGGTGCTCAAGATACCTGGGTCTGTTGGCGGC 1920
 Qy 641 ValIleuAlaIleAlaValAlaValTyrcyLeuSerThrGlyCyValValIleValGlyArg 660
 Db 1921 GTTCTGCTGCTTTGGCGCGGTATGCTTATCCACAGGCTGCGGTGCTCATATGATGAGG 1980
 Qy 661 IleValIleuSerGlyLyProAlaIleIleProAAsArgGluValIleuTyArgGluPhe 680
 Db 1981 ATTGTCTTGTCCGGAAGCGCGCAATCATACCGACAGGGAAGTCTCTACCGGGAATTTC 2040
 Qy 681 AspGluMetGluGluCys 686
 Db 2041 GATGAAATGGAAAGTGC 2058
 RESULT 2
 AX467113
 LOCUS AX467113 2061 bp DNA linear PAT 17-JUL-2002
 DEFINITION Sequence 1 from Patent WO0214362.
 ACCESSION AX467113
 VERSION AX467113.1 GI:21900425
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1
 AUTHORS Sallberg, M.
 TITLE A hepatitis C virus non-structural ns3/4a fusion gene
 JOURNAL Patent: WO 0214362-A 1 21-FEB-2002;
 TRIPEP AB (SE)
 FEATURES
 source 1..2061
 Location/Qualifiers
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Hepatitis C virus NS3/4A coding region"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,41e-203 Length: 2061
 Score: 3618.00 Matches: 686
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-530-591-2 (1-686) x AX467113 (1-2061)
 Qy 1 MeAlaProIleThraIaTyraGlnGlnThraArgGlyLeuLeuGlyCyAlaIleIleThr 20
 Db 1 ATGGCGCTATCAAGCGCTATGCCAGACAGCAAGGGGCTTTGGATGCTATATACCC 60
 Qy 21 SerIleuThrGlyAAspGlyAAsngluValGluGluValGlnIleValIleSerThra 40
 Db 61 AGCTTGACCGGCGCGGACAAACACAGGTGAGGGGTGAGTTCAAGTCTGTCAACTGCT 120
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAAsnglyValCysTrpThraValTyrlleGlyAla 60
 Db 121 GCCAGACTTCTTGGCACTGCACTTACCGGGGTGTGTGACTGTCTACCAATGAGACC 180
 Qy 61 GlyThraArgThrIleAlaSerProGlySerProValIleGlnMetTyrlThraAsnValAsp 80

Db 181 GGAACAGAGCAATTGCTACCTAAGGCTCTGTTATTCAGATGTACCAATGTGAC 240
 Qy 81 GlnAAspleuValGlyTrpProAlaProGlnGlyAlaAArgSerIleuThrProCyTrhCys 100
 Db 241 CAGACCTCGTAGGCTGTGGCCGCTCCCAAGGTGCCGCTCATTAACCAATGCACTTGG 300
 Qy 101 GlySerSerAAspleuTyrlleuValThraArgHisAlaAAspValIleProValAArgArg 120
 Db 301 GGCTCTCGGACTTAACTTACCTGTGACAGAGCAAGCCGAGATGTACTTCTGTGCCCGCAGG 360
 Qy 121 GlyAAsnglyAArgGlySerIleuSerProAArgProIleSerTyrlleuLyGlySerSer 140
 Db 361 GGTAAGGCAAGGGGAGCGCTGCTTGGCCCGCTTATCTTATCTTAAAGGCTCTTCC 420
 Qy 141 GlyGlyProleuLeuCyAspAlaGlyHisAlaValGlyIlePheAArgAlaAlaValCys 160
 Db 421 GAGAGCCCTCTGTGTGTGCCCCGAGAGACATGCCGTAGGCAATTCAGAGCCGGGTATGC 480
 Qy 161 ThraArgGlyValAlaAlaValAlaAAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 481 ACCCGTAGAGTGTGCTAAGGCGGTGACTTCACTCCGTTAGAGCTTAAAGACAAACATG 540
 Qy 181 ArgSerProValAAspSerAAspAAsnSerSerProProAlaValAProGlnSerTyrlGlnVal 200
 Db 541 AGGTCCCGGTGTCTCAGCAACTCTCCCAACAGAGTGCCTCAAGATG 600
 Qy 201 AlaHisIleuHisAlaIleProThrGlySerGlyLySerSerThryValProAlaAlaTyrla 220
 Db 601 GCCCACTGCTGATCTCTCCACCGGACGGTAAAGACCAAGTCCCGCGGCAATACGA 660
 Qy 221 AlaGlnGlyTyrllyValIleuValIleuAAsnProSerValAlaIleThrMetGlyPheGly 240
 Db 661 GCTCAGGGGCTACAAAGGTGTGGTGTCAACCCCTCGTTGCTCAACAATGGGCTTTGGT 720
 Qy 241 AlaTyrlMetSerTyrlsalAHisGlyIleAAspProAAsnIleArgThrGlyValAArgThrIle 260
 Db 721 GCTTACATGTCCAAAGGCCCAATGGATGTATCTTAACATCAGACTGGGTGAGACAAATT 780
 Qy 261 ThrThrGlySerProIleThrTyrlSerThrTyrlGlyLyPheLeuAlaAAspGlyGlyCys 280
 Db 781 ACTACTGGAGCGCCGATACAGTATTCACCTTACCGGCAAGTCTTGTCCAGCGCGGTGT 840
 Qy 281 SerGlyGlyAlaTyrlAAspIleIleIleCyAAspGluCyHisSerThraAAlaThrSer 300
 Db 841 TCAGGGGGGTCTTATGACATTAATTTGTGAGAGAGGCACTCCACGATGCAATCC 900
 Qy 301 IleLeuGlyTyrlleGlyThrValIleuAAspGlnAlaGluThraIleGlyAlaArgGluThrVal 320
 Db 901 ATCTTGGCAATTGGCACTGTCTTGTACCAAGCAGAGACCGCGGGCGAGACTGCTGT 960
 Qy 321 LeuAlaThraIleThraProProGlySerValThrValProHisProAAsnIleGluGluVal 340
 Db 961 CTCGCACCGCTACCCCTCCGGGCTCCGTCACGTGTGCCCATCTTAATCGAGAGGTT 1020
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrlGlyLyAlaIleProleuGluAlaIle 360
 Db 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTATGGCAAGGCTATTCCTTGAAGAATT 1080
 Qy 361 LyGlyGlyAArgHisIleuIlePheCyHisSerTyrllyLyCyAAspGluLeuAlaAla 380
 Db 1081 AAGGGGGGAGACATCTCTTCTGTCCACTCAAAAGAGAGCAGCGAGCTCGCGCA 1140
 Qy 381 LyLeuValAlaIleuGlyValAAsnAlaValAlaTyrlTyrlArgGlyLeuAAspValSerVal 400
 Db 1141 AAATGTGTGCGTTGGCGCTCAATGCGGTGACTTACACGGCGCTTGAATGTGTGCTC 1200
 Qy 401 IleProThrSerGlyAAspValValAlaValAlaIleThraAAspAlaLeuMetThrGlyPheThr 420
 Db 1201 ATCCGACCAAGTGTGACGTTGTGCTGTGGCAACTGAGCCCTTATGACCGGCTTATCC 1260
 Qy 421 GlyAAspPheAAspSerValIleAAspCyAAsnThrCysValThrGlnThrValAAspPheSer 440
 Db 1261 GCGCACTTCAATTCGATGTAGTCACTCAACACGTGTGTCAACCAAGTGTGACTTACG 1320

QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 1321 CTTGACCTTACCTTACCATTTGAGACATACAGCTTCCCGAGATGCTGCTCCCGACT 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLeuProGlyIleTyrArgPheValAlaProGly 480
 Db 1381 CAAAGTCGGGGTACGATCGGACAGAGGAGAACCGAGCATCTACAGATTGTGGACCGGG 1440
 QY 481 GluArgProSerGlyMetPheAspSerValLeuGlyGlyCysTyrAspAlaGlyCys 500
 Db 1441 GAGGCTCTTCTGCAATGTTTGACTCGTCTGCTCTGCAAGTGTATGACGGGGTGT 1500
 QY 501 AlaTyrTyrGlyLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 1501 GCTTGGATAGAGTTAGCCCGCCGAGACCAAGTATAGGCTACAGAGCATACATGACACC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGlyValPheThrGlyLeu 540
 Db 1561 CCGGGACTTCCCGTGTGCAAGACCATCTTGAAATTTGGAGGGCGTCTTTACGGGTCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrTyrGlnSerGlyGluAsnLeuProTyr 560
 Db 1621 ACCCAATAGAGCGCCACTTCTATCCAGCAAGACAGAGTGGGAAAACTTCCCTAT 1680
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
 Db 1681 CTGTAGAGTACCAAGCCACCGTGTGCTAGACTCAAGCCCTTCCCGTGTGGAC 1740
 QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGAAGTGTGTTATCCGTCTCAAGCCACCTCCATGGGCAACCTCTGCTA 1800
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 1801 TATAGACGGGCGCTGTCAGAAATGAATCACTCGACGACCCAGTCACCAAGTATATTC 1860
 QY 621 MetTyrCysMetSerAlaAspLeuGluValAlaThrSerThrTyrValLeuValGlyGly 640
 Db 1861 ATGACATGATATGTCGGCTGACCTGAGGTGCTCAAGACTACCTGGGTGCTGTTGGCGGC 1920
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 1921 GTTCTGGCTGTTGGCGGATATGCTATCCAGGCTCGTGTGTCATGTAGTGG 1980
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 1981 ATTGTCTGTCCGAAAGCCGCAATCAACCCGAGAGGAGTCTTACCGGAGTTC 2040
 QY 681 AspGluMetGluGluCys 686
 Db 2041 GATGAATGGAAGAGTGC 2058
 RESULT 3
 LOCUS 106434 5360 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 48 from Patent EP 0318216.
 ACCESSION 106434
 VERSION 106434.1 GI:590311
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES 1 (bases 1 to 5360)
 AUTHORS Houghton M., Choo O.-T., and Kuo G.
 TITLE Nanby diagnostics and vaccines
 JOURNAL Patent: EP 0318216-A1 48 31-MAY-1989;
 FEATURES Location/Qualifiers
 1..5360
 /organism="Unknown"
 /mol_type="unassigned DNA"

Alignment Scores:

Pred. No.: 1,74e-200 Length: 5360
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x 106434 (1-5360)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 930 CGGCGCCCATACGCGCGTACCGCCAGACAGACAGGGGCTCTTGGGTGATATAC 989
 QY 21 SerLeuThrGlyArgAspLysAsnGluValGluGlyValGlnIleValSerThrAla 40
 Db 990 AGCTTACTGCGCGGACAAACCAAGTGGAGGTGAGGTCAAGATTGTGCACTGCT 1049
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyVal 60
 Db 1050 GCCCAACCTTCTGTGCAACGTGCAATCAATGGGGTGTGTGACTGTCAACGAGGCTC 1109
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1110 GGAACGAGACCATCCGCTACCCAGAGGCTCTGTATCCAGATGATATACCAATGTAGC 1169
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1170 CAAGACTTGTGGGCTGCGCCGCTCGCAAGTAGCGCTCATTTGACACCTTGCACTTGC 1229
 QY 101 GlySerSerAspLeuTyrIleValIleThrArgHisAlaAspValIleProValArgArgArg 120
 Db 1230 GGCTCTCGACCTTTACTGTGTACAGAGACAGCGCGAGTATTCCTCGTGGCGGCGG 1289
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleLysGlySerSer 140
 Db 1290 GGTGATAGACGGGGACGCTGTGTGCGCCGCCATTTCTCACTTGAAGGCTTCCTGC 1349
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 1350 GGGGCTCCGCTGTGTGGCCCGCGGGGACGCGGTGGGCAATTTAGGGCCGCGGTGTC 1409
 QY 161 ThrArgGlyValAlaValAspAlaValAspPheIleProValGlySerLeuGluThrThrMet 180
 Db 1410 ACCGTGAGTGTCTAAGGGGAGGAGTATTCCTGTGTGAGAGACTTACAGCAACAG 1469
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 1470 AGGTCCCGGTGTTCAAGATATCTCTCCACAGTAGTGCCGACAGAGCTTCCAGGTG 1529
 QY 201 AlaHisLeuHisAlaProThrArgLysSerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 1530 GCTCAGCTCCATGCTCCACAGGACGCGCAAAAGACCAAGGTCCGGCTGATATGCA 1589
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 1590 GCTCAGGAGCTATTAAGTGTAGTACTCAACCTCTGTGTGTGCAACATCGGGCTTTGGT 1649
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 1650 GCTTACATGTCAAGGCTCATGGAGTCATCTTACATCGACGCGGGGAGAACAAATT 1709
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 1710 ACCACTGGACGCCCATACAGTACTCCACTTACGAGAGTCTTGTCCGACGCGGGGTGC 1769
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 1770 TCGGGGGGCGCTTATGACATTAATATTGTGAGAGTGCACATCCAGGATGCACTCC 1829
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValaArgLeuThrVal 320
 Db 1830 ATCTTGGGCAATCGGCACTGTCTTGAACAGAGAGACTCGGGGGGAGACTGTTGTG 1889

QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 1890 CTCGCCACCGCCACCCCTCCGGGCTCCGTCATGTCGCCATCCAAATCGAGAGGTT 1949
 QY 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyValAlaIleProLeuGluAlaIle 360
 DB 1950 GCTCTGTCACCAACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAGATATC 2009
 QY 361 LysGlyValArgHisLeuIlePheCysHisSerIleValIleCysAspGluLeuAlaAla 380
 DB 2010 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAGAGAGAGTGCAGACATCTCGCGCA 2069
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaIleThrValArgGlyLeuAspValSerVal 400
 DB 2070 AAGCTGGTGCATTTGGGCATCAATGCGGTGGCTACTACCGCGGTCTTGAAGTGGCTG 2129
 QY 401 IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuThrGlyPheThr 420
 DB 2130 ATCCCGACCAAGCGCGATGTTGTCGTGGCAACCATGCTCATGACCGGCTATACC 2189
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 2190 GGGGACTTGCATCTGGATGATGACTGCATATCGTGTCAACCAAGACATCGAATTCAGC 2249
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 2250 CTTCACCTTACCTTTCACCATTTAGACATACAGCTCCCGACAGATGTCGTCCCGACT 2309
 QY 461 GlnArgArgGlyValArgThrGlyValArgGlyLysProGlyIleThrArgPheValAlaProGly 480
 DB 2310 CAACGTCGGGGGACAGCTGGCAGGGGAGACCAAGCATCTACAGATTTGTGGCACCGGG 2369
 QY 481 GlnArgProSerGlyMetPheAspSerValLeuGluGlyCysThrAspAlaGlyCys 500
 DB 2370 GAGGGCCCCCTCCGCGATGTCGACTCGTCCGCTCTGTAGTGTATGACGAGGCTGT 2429
 QY 501 AlaTrpTrpGluLeuThrProAlaGluThrThrValArgLeuArgAlaIleTrpMetAsnThr 520
 DB 2430 GCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2489
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 DB 2490 CCGGGGCTTCCCGGTGTCAGACATCTTGAATTTGGAGGGGCTTTTACAGGGGCTC 2549
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyValLeuLeuProTrp 560
 DB 2550 ACTCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2609
 QY 561 LeuValAlaIleThrGlnAlaThrValCysAlaArgAlaGlnAlaProProPheSerTrpAsp 580
 DB 2610 CTGGTAGCTTACCAAGCCACCGTGTGCTGAGGCTTAAAGCCCTCCCTCCATGCTGGGAC 2669
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 2670 CAGATGTGAAGTGTGATTCGCTCAAGCCACCTTCATGGGCAACACCCCTGCTA 2729
 QY 601 TyrArgLeuGlyValAlaGlnAsnGluValIleLeuThrHisPheValIleThrIle 620
 DB 2730 TACAGACTGGGCGGTGTTCAAGATGAAATCAACCTTCAGCACCAAGTCAACCAATATATC 2789
 QY 621 MetThrCysMetSerAlaAspLeuGluValIleThrSerThrTrpValIleValIleGly 640
 DB 2790 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2849
 QY 641 ValLeuAlaIleAlaValAlaIleThrCysLeuSerThrGlyCysValValIleValIleGly 660
 DB 2850 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2909
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuIleThrArgGluPhe 680
 DB 2910 GTGCTTCTTCTCGGGAAGCGGCAATCATCTGACAGGAGAAATGCTCTACGAGATTC 2969

QY 681 AspGluMetGluGluCys 686
 DB 2970 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2987
 RESULT 4
 LOCUS 109328 5360 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 8 from Patent WO 8904669.
 ACCESSION 109328
 VERSION 109328.1 GI:587963
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 5360)
 AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
 JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;
 FEATURES
 source Location/Qualifiers
 1..5360
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,74e-200 Length: 5360
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x 109328 (1-5360)
 QY 1 MetAlaProIleThrAlaThrAlaGlnIleThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 930 CTGGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 990 AGCTTAACCTGGCGGACCAAAACCAAGTGAAGGAGTCCAGATTGTGTCAACTGCT 1049
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleHisGlyAla 60
 DB 1050 GCCCAACCTTCTCGGCAAGTGCATCAATGAGGCTGTGCTGACATCTTACACCGGGGCT 1109
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTrpThrAsnValAsp 80
 DB 1110 GGAACAGAGACCATCGCGTACCAAGGCTCTGTATCCAGATGTATCAATGTATGAC 1169
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 1170 CAGAGCTTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1229
 QY 101 GlySerSerAspLeuTrpLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 1230 GGCTCTCGGACCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
 QY 121 GlyAspGlyValArgGlySerLeuLeuSerProArgProIleSerTrpLeuLysGlySerSer 140
 DB 1290 GGTGATAGAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1349
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValAlaGlyIlePheArgAlaAlaValCys 160
 DB 1350 GGGGGTCCGCTGTGTGTCCTCCCGGGGACGCGCGTGGGCAATTTTGGGCGCGGTGTGC 1409
 QY 161 ThrArgGlyValAlaValAlaValAlaAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 1410 ACCCGTGAAGTGGCTAAGCGGAGCTTATCCCTGTGAGAGAACTTACAGCAACCATG 1469
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTrpGlnVal 200
 DB 1470 AGGTCCCGGCTGTTCAGGATTAATCTCTTCCACCAAGTAGTGCCTCCAGAGTTCAGG 1529

934
 1213
 1934
 273
 934-1213 > 100

QY 201 AAlHsiLeuHsAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 1530 GCTCACTCCATGCTCCACAGGAGCGGCAAAAGCAACCAAGGTCCCGGTGATATGCA 1589
 QY 221 AAlGlnGlyTyrLysValLeuValLeuAspProSerValAlaAlaThrMetGlyPheGly 240
 DB 1590 GCTCAGGAGCTAATAGGCTAGTACTACACCCCTCTGTTCTGTCACACCTGGGCTTTGGT 1649
 QY 241 AAlaTyrMetSerLysAlaHsiGlyLysLeuAspProAsnLysLeuThrGlyValArgThrTle 260
 DB 1650 GCTTACATGTCACAGGCTCATGGGATGATCTTAATCATCAGACCGGGGAGAAACAAT 1709
 QY 261 ThrThrGlySerProLysThrLysThrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 1710 ACCACTGGCAGCCCACTACAGTACTCCACCTACGCAAGTCTCTGCGAGCGGGGTGC 1769
 QY 281 SerGlyGlyAlaTyrAspLysLeuLysCysAspGlyLysSerThrAspAlaThrSer 300
 DB 1770 TCGGGGGGGCGCTTATGACATTAATTTGTGACAGAGTCCACATCCAGATGCCAATCC 1829
 QY 301 LLeuGlyLysGlyThrValLeuAspGlnAlaGluThrAlaGlyValArgLeuThrVal 320
 DB 1830 ATCTGGGCACTGGCACTGCTTACCAAGCAGAGACTGGGGGCGAGACTGTTGTG 1889
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnLysGluVal 340
 DB 1890 CTCGCCACCGCCACCTCCGCGCTCCGCTCCTGCTACCTGCCCCCAATCGAGAGGTT 1949
 QY 341 AAlaSerSerThrThrGlyGluLysProPheTyrGlyLysAlaLysProLeuGluVal 360
 DB 1950 GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGGTATCCCTCGCAAGTATC 2009
 QY 361 LysGlyGlyValArgHsLeuLysPheCysHsSerLysLysCysAspGluLeuAlaAla 380
 DB 2010 AAGGGGGGAGACATCTCATCTCTCTGCTCAATCAAGAGAGTCCGAGAACTGCGCA 2069
 QY 381 LysLeuValAlaLeuGlyValAspAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 2070 AAGCTGTGCGATTGGGCACTAAGTCCGTGGCTTACACCGGCTTGAAGTCTGCTC 2129
 QY 401 LLeuProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 2130 ATCCGACAGAGGGGAGTGTGCTGCTGCGCAACCGATGCCCCATGACCGGCTATACC 2189
 QY 421 GlyAspPheAspSerValLysAspCysAsnThrCysValLysThrValAspPheSer 440
 DB 2190 GCGCACTTCGACTCGGTATGACATGCAATACGTGTGCACCGACAGCTGATTTACG 2249
 QY 441 LeuAspProThrPheThrLysGluThrLysLeuProGlnAspAlaValSerArgThr 460
 DB 2250 CTGACCTTACCTTACCATTTAGACATCACTGCTCCCGAGATGCTGCTCCGCACT 2309
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyLysTyrArgPheValAlaProGly 480
 DB 2310 CAACGTCGGGGAGAGACTGCGAGGGGAGAGCGGCACTACAGATTGTGTGCAACCGGG 2369
 QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGlyTyrAspAlaGlyCys 500
 DB 2370 GAGGCCCTCCCGGATGTGACTGCTGCTCTCTGTAGTGTCTTACAGCAGGCTGT 2429
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 2430 GCTTGGTATGAGCTACCGCGCGGACCTACAGTTAGCTACGAGCTATGAAACACC 2489
 QY 521 ProGlyLeuProValCysGlnAspHsLeuGluPheTyrGluGlyValPheThrGlyLeu 540
 DB 2490 CCGGGGCTTCCCGGTGTGCGAGACCATCTTGAATTTGGAGAGGGGTCTTACAGGCTTC 2549
 QY 541 ThrHsLysAspAlaHsPheLeuSerGlnThrLysGlnSerGlyGlyAsnLeuProTyr 560
 DB 2550 ACTCATATATGATGCCACTTCTTATCCGACAAAGCAGAGTGGGAGAACTTCTTAC 2609
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrPasp 580

DB 2610 CTGTAGCGTACCAAGCCACCGGTGTGGGCTAGGGCTCAAGCCCTCCCATCGGGAGC 2669
 QY 581 GlnMetTyrLysCysLeuLysLeuArgLeuLysProThrLeuHsGlyProThrProLeuLeu 600
 DB 2670 CAGATGTGAAGTGTATTTATTCCTCAAGCCCACTCCCATGGGCAACACCTTGCTA 2729
 QY 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrLysTyrTle 620
 DB 2730 TACAGCTGGGGGCTGTTCAGATGAATATCACTCGAGCCAGTCCATGCAACCAATATC 2789
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrPValLeuValGlyGly 640
 DB 2790 ATGACATGATGTCCGGCGCACTGGAGGTGTCACAGACACCTGGGTCTCGTGGCGG 2849
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValLysValArg 660
 DB 2850 GTCTGGCTGCTTGTGGCGCGTATTTGCTGTAAACAGGCTGCGTGTCTATAGTGGCGG 2909
 QY 661 LLeuValLeuSerGlyLysProAlaLysLeuProAspArgGluValLeuTyrArgGluPhe 680
 DB 2910 GTGCTTGTTCGGGAAAGCGGCAATCACTGACAGGAAAGTCTTACCGAGATTC 2969
 QY 681 AspGluMetGluGluCys 686
 DB 2970 GATGAGATGGAAGATGC 2987
 RESULT 5
 LOCUS 106440 6785 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 54 from Patent EP 0318216.
 ACCESSION 106440
 VERSION 106440.1 GI:590312
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 6785)
 AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
 TITLE NAbv diagnostics and vaccines
 JOURNAL Patent: EP 0318216-A1 54 31-MAY-1989;
 FEATURES
 source 1..6785
 /organism="Unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,32e-200 Length: 6785
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x 106440 (1-6785)
 QY 1 MetalaproteinHsAlaTyrAlaGlnGlnThrArgGlyLeuGluGlyCysLysLysLeu 20
 DB 1203 CTGGCGGCCATCAAGGGTATGCGCCAGACAGAAAGGGGCTCTTACGGTGCATATACCC 1262
 QY 21 SerLeuThrGlyValArgAspLysAsnGlnValGluGlyGluValGlnLysLeuThrAla 40
 DB 1263 AGCTAATCTGGCGGCGCAAAAACAAAGTGGAGGGTGAAGGTCCAGANTGTGTCACTGCT 1322
 QY 41 AAlaGlnThrPheLeuAlaThrCysLysLeuGlyValCysTyrThrThrValTyrHsGlyAla 60
 DB 1323 GCCCAACCTTCTCCGCAACGTGATCAATATGGGTGTGCTGACATGCTACACAGGGGCC 1382
 QY 61 GlyThrArgThrLysLeuAspProLysGlyProValLysGlnMetTyrThrAsnValAsp 80
 DB 1383 GGAACGAGGACATCGGTCAACCAAGGGTCTGTATCAGATGTATACCAATGTATGAC 1442

QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyValAargSerLeuThrProCysThrCys 100
 Db 1443 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTAGCGCTCATTTGACACCTGGACCTTGC 1502
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 Db 1503 GGGCTCCCGACCTTTTACTGTGTACAGAGGACCGCATGTCTTCCCGCGCGCGG 1562
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlySerSer 140
 Db 1563 GGTGATAGCAGGGGACGCTGCTCTCCGCGGCCCATTTCTTCACTTGAAGGCTCTTCG 1622
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 1623 GGGGGTCCGCTGTGTGCTCCCGCGGGGACGCGGTGGCATATTAGGCGCGGTGTGC 1662
 QY 161 ThrArgGlyValAlaValAspPheIleProValGlySerLeuGlyThrThrMet 180
 Db 1683 ACCGTGTAGTGTGGTGAAGCGGTGTGACTTATCCCTGTGAGAACCTTAGAACCAACATG 1742
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 1743 AGGTCCCGGTGTTACGATTAAGTCTCTTCCACCAAGTAGTCCCGACGCTTCAGAGTG 1802
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 1803 GCTACCTTCATGCTCCACAGGACGCGGCAAGAACCAAGTCCCGGTGATAGCA 1862
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 1863 GCTCAGGGCTATAGGCTAGTACTCAACCCCTGTGTGTGTCGACACGCGGCTTGTGT 1922
 QY 241 AlaTyrMetSerLysValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 1923 GCTTACATGTCACAGGCTCATGGATGATCTTACATCAGACCGGGGTGAGAACAAATT 1982
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 1983 ACCACGTGAGCCCATCAGTACTCCACCTAGCGGAAGTTCCTTGCAGCGGGGTGTGC 2042
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
 Db 2043 TCGGGGGGCGCTTATGACATATATTGTGACAGAGGCCACTCCAGGATGCCACATCC 2102
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlyThrAlaGlyValAargLeuThrVal 320
 Db 2103 ATCTTGGGCAATCGGACATGCTCTTGACCAAGACAGACTGGGGGGGAGACTGTGTGTG 2162
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 2163 CTGGCCACCGGCACCTCCGGGGCTCGCTCATCTGTGCCCATGCCAACATCGAGAGGTT 2222
 QY 341 AlaLeuSerThrThrGlyGlyIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2223 GCTCTGTGCCACCAACGAGAGATCTCTTTTACCGGCAAGGCTATCCCTCGAAGTATC 2282
 QY 361 LysGlyGlyIleArgHisLeuIlePheCysHisSerTyrLysValCysAspGlyLeuAlaAla 380
 Db 2283 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAGAGAGTCCGACGAACTGCCGCA 2342
 QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 2343 AAGGTGTGTCATTTGGGACATCAAGCGGTGCTGCTGACCGCGTCTTGAAGTCTCGTC 2402
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2403 ATCCCGACGACGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2462
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 2463 GGGACATTTCACTCGGATGATGACTGCAATACGTGTGTGTGTGTGTGTGTGTGTGTGT 2522
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460

Db 2523 CTTCACCTTACCTTCAACATTTAGACATCAACCTCCCGACGATGCTGTCCGCACT 2582
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 2583 CAACGTGGGGGAGACATGCGAGGGGAGACCGGACATTTACAGATTTGTGGACCGGG 2642
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTyrAspAlaGlyCys 500
 Db 2643 GAGGGCCCTCCGCGCATGTTTGCATCGTCCGTCTGTGATGCTATGACGAGGCTGT 2702
 QY 501 AlaTyrTyrGlyLeuThrProAlaGlyThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 2703 GCTTGATATAGCTCAACGCGCGGACGACTACAGTTAGGCTACGAGGCTACAGAACCC 2762
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGlnGlyValPheThrGlyLeu 540
 Db 2763 CCGGGGCTTCCCGTGTGCGCAGACCATTTGAATTTGGAGGGCGCTTTTACAGGCTTC 2822
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 2823 ACTCATATAGATGCCACCTTCTATCCACAGAACAGAGTGGGAGAACCTTCTTAC 2882
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
 Db 2883 CTGGTAGCGTACCAAGCCACCGTGTGCTAGGGCTCAAGCCCTCCCATGCTGGGAC 2942
 QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 2943 CAGATGTGAAGTGTGTGATTTGCTCTCAAGCCCACTTCATGAGCCCAACCCCTGTCTA 3002
 QY 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 3003 TACAGACTGGCGCTGTTCAGATGAATATACCTTACAGCAACCAAGTACCAATATCATC 3062
 QY 621 MetThrCysMetSerAlaAspLeuGlyValValThrSerThrTyrValLeuValGlyGly 640
 Db 3063 ATGACATGACATGTGCGCGCCGACCTGAGGTGTGCACAGCACCGGGGTGCTGTGGGGCG 3122
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 3123 GTCTGTGCTGCTTTGGCGCGGATGTCTGTCTGTCAACAGGCTGCGTGTGATAGTGGGACG 3182
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGlyValLeuTyrArgGluPhe 680
 Db 3183 GTCTGTGCTGCGGAGACCGGACATCATCTGACAGGAAATCTCTTACCGAGATTC 3242
 QY 681 AspGluMetGluGluCys 686
 Db 3243 GATGAGATGGAAGATGC 3260

RESULT 6
 LOCUS 109329 6785 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 10 from Patent WO 8904669.
 ACCESSION 109329
 VERSION 109329.1 GI:587964
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES 1 (bases 1 to 6785)
 Houghton, M., Choo, O.-K. and Kuo, G.
 Patent: WO 8904669-A 10 01-JUN-1989.
 JOURNAL Location/Qualifiers
 FEATURES
 source /mol_type="unknown"
 ORIGIN
 Alignment Scores: 2.32e-200 Length: 6785
 Pred. No.: 3574.00 Matches: 672
 Score:

Percent Similarity: 99.42%
 Best Local Similarity: 97.96%
 Query Match: 98.78%
 DB: 6

Conservative: 10
 Mismatches: 4
 Indels: 0
 Gaps: 0

US-09-930-591-2 (1-686) x 109329 (1-6785)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1203 CTGCGCCCATACAGGGGTACGCCAGACAGAGGGGCTCTCAGGGGTCAATATCACC 1262
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
 DB 1263 AGCTTAACCTGCGCGGCAAAAACCAAGAGGTGAGGTCCAGATTGTGTCAACGCT 1322
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrGlyVal 60
 DB 1323 GCCCAACCTTCTCGGCAACGTGCATCATGCGGTGTGTGACTGTCTACACGGGGCC 1382
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 1383 GGAACGAGAACATCGGTGTCACCCAGAGGTCTGTCTACAGATGTATACCAATGTAGAC 1442
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 1443 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGGTAGCCGCTCATTTGACACCTGCACTTC 1502
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 DB 1503 GGCTCCTCGGACCTTACCTGTGTGTCAGAGGACCGCATGTATCCCGTGGCCGGCGG 1562
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 1563 GGTATATGCAAGGGGCACTGTCTGTGCGCCGCGCATTTCTTCTTGAAGGCTCTTCG 1622
 QY 141 GlyLysProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 1623 GGGGGTCCGCTGTGTGTGCGCGGGGACGCGTGGGCATATTAGGGCGCGGCTGTGC 1682
 QY 161 ThrArgGlyValAlaIleValAlaValAspPheIleProValGlySerLeuGlnThrMet 180
 DB 1683 ACCGTGTAGTGGCTAAGCGGTGAGCTTATCCCTGTGAGAACCTAAGAACCAACATG 1742
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 1743 AGGTCCCGGTGTTTCAAGATTAATCTCTTCCACCAAGTAGTCCCGAGCTTCCAGTG 1802
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 1803 GCTACACCTCCAGTCCCAAGGACGAGCGGCAAAAGCACCAAGGTCCGGCTGATATGCA 1862
 QY 221 AlaGlnGlyTyrValValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 1863 GCTAAGGGCTTAAAGGTCTAGTACTAACCCCTCTGTCTGTGAAACACTGGGCTTTGGT 1922
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 1923 GCTTACATGTCCAAAGCTCATGGATGCAATCTTAACATCAGAACCGGGGTGAAACAAAT 1982
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 1983 ACCACTGGCAGCCCATCATCAGTACCTACCTACGGCAAGTTCCTTCCGCAAGCGGGGCG 2042
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
 DB 2043 TCGGGGGGCGCTTATGACATATATTTGTGACGAGTCCACTCCAGGATGCCACATCC 2102
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 2103 ATCTTGGGATCGGACATGTCTTGTACCAAGCAAGACCTGCGGGGCGAGACTGTTCTG 2162
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340

DB 2163 CTCGCCACGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCAAATCGAGAGGTT 2222
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGlnAlaIle 360
 DB 2223 GCTGTGTCCACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCAAAGTATC 2282
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaIle 380
 DB 2283 AAGGGGGGCAAGCATCTCATCTTGTGTATTCAAAGAAAGATGCGACAACTCGCGGCA 2342
 QY 381 LysLeuValAlaLeuGlyValAlaAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 2343 AAGCTGTGTGCAATGGGCATCATGCGGGCTCTACCGGGTCTTGACGTGTCCGTC 2402
 QY 401 IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 2403 ATCCGACAGAGGCGCATTTGTCTGTGCGCAACGATGCCCTCATGACCGGCTATCC 2462
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrValAspPheSer 440
 DB 2463 GCGGACTTGCATCGGTGATGACGTGCAATACGTGTCAACCAAGCATGCTATTCACG 2522
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 2523 CTTGACCTTACCTTCAACATTTGACATACAGCTCCCAAGATGTGTCTCCGACT 2582
 QY 461 GlnArgArgGlyValArgThrGlyArgLysProGlyIleTyrArgPheValAlaProGly 480
 DB 2583 CAACGTGGGGGAGAGCTGGAGGGGAGAGCCAGATCTTACGATTTGTGGACCGGGG 2642
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 DB 2643 GAGCGCCCTCCGCGCATGTGTGACCTGCTCCGTCTGTGAGTGTATGACGAGCTGT 2702
 QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 2703 GCTTGTATGAGCTCAAGCCCGGAGACTACGTTAAGGTCTACAGAGCTATCAAGAC 2762
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 DB 2763 CCGGGGCTCCGCTGTGCGCAGACATCTTGAATTTGGAGGGGCTTTTACAGGCTTC 2822
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysLeuLeuProTyr 560
 DB 2823 ACTCATATAGATCCCACTTCTATCCAGACAAACAGAGTGGGAGAACCTTCTTAC 2882
 QY 561 LeuValAlaIleProGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
 DB 2883 CTGTAGCCGTACCAAGCCACCGTGTGCTAGGGCTCAAGCCCTCCCATCTGTGGAC 2942
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 2943 CAAGTGTGAAAGTGTGATTCGCCCTCAAGCCCAACCTCATGGGCAACCCCTGTCTG 3002
 QY 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 DB 3003 TACAGACTGGGCGCTGTTCAGATGAATCAACCTCAGCAGCAGCAGCAGCAAAATCATC 3062
 QY 621 MetThrCysMetSerLysLeuGluGluValIleValThrSerThrTrpValLeuValGlyGly 640
 DB 3063 ATGACATGATGTGCGCGGACCTGAGAGTGTACAGACCTGGGTGCTCTGTGGCGGC 3122
 QY 641 ValLeuAlaIleLeuAlaIleTyrCysLeuSerThrGlyCysValIleValIleValArg 660
 DB 3123 GTCTGTGCTGTCTTGGCGCGGTATTCGTCTGTCAACAGGCTGCGTGGCATATAGTGGGAG 3182
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 DB 3183 GTCGTCTTGTCCGGGAAGCGGCAATCATCTGACAGGAGAGTCTCTTACGAGAGTTC 3242
 QY 681 AspGluMetGluGluCys 686
 DB 3243 GATGATGATGAGAGTGTCC 3260

RESULT 7
LOCUS ARI18696 7310 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 74 from patent US 6150087.
ACCESSION ARI18696
VERSION ARI18696.1 GI:14100606
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 7310)
AUTHORS Chien, D. Y.
TITLE NANBV diagnostics and vaccines
JOURNAL Patent: US 6150087-A 74 21-NOV-2000;
FEATURES Location/Qualifiers
source 1..7310
/organism="unknown"
/mol_type="unassigned DNA"

Alignment Scores:
Pred. No.: 2.54e-200 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x ARI18696 (1-7310)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCGCCCATTCACGGGCTACGCCACACAGAGGGGCTCTTCAAGGTGCTATATACC 1787
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGlnGlyGluValGlnIleValSerThrAla 40
Db 1788 AGGCTAACTGGCCGGGACAAACCAAGTGGAGGTGAGTCCAGATTGTCAACTGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1848 GCCCAAACTTCTCGGACAGTGCATCAATGGGTGTGCTGACTGTCTACCAAGGGGCC 1907
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGAACGAGACCATCGGTCTACCAAGGTCCTCTCATCCAGATGTATACCAATGTAGAC 1967
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyValAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAGACCTTGCGGCTGGCCGCTCCGCAAGTACCGCTCATTTGACACCCCTGCATTGC 2027
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 2028 GGCTCTCGGACCTTTCACCTGTGTCACAGAGCACCGCATGTCCCGTGGCCGCGCG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 2088 GGTGATAGCAGGGGACAGCTGCTGTCCGCCGCCCATTTCTTCACTTGAAGAGCTCTTCG 2147
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 2148 GGGGGTCCGCTGTGTGCCCCGCGGGGACGCGTGGGCATATTAGGGCGCGGTGTGC 2207
QY 161 ThrArgGlyValAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCGTTGAGTGTAGCGGTGAGCTTATCCCTGTGAGAACCTTAGAGACAACATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGGTCCCCGGTGTTCACGATTACTCTTCCACCAAGTAGTGCCCAAGGCTTCAAGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220

Db 2328 GCTCACTCCATGCTCCCAAGGACGGGCAAAAGACCAAGGTCCCGGCTGCATATGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATTAAGGTGTAGTACTCAACCCCTGTGTGCTGCACACACTGGGCTTTGGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCCAAGGCTCATGGATCGATCTTCAATCATCAGACCGGGGTGAAACAATT 2507
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCACTGGCAGCCCATTCAGTACTCCACTTACGGCAGAGTCTTGTGCGAGCGGGGTGC 2567
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGGCGCTTATGACATATATTTTGAAGAGTCCACTCCAGATGCCACATCC 2627
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTGGGCATCGGACGTGCTTGACCAAGCAGACATCGGGGGGAGACTGTGTGTG 2687
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTCGCCACCGCCACCCCTCGGGGCTCGTCACGTGTGCCATCCCAACATCGAGAGGTT 2747
QY 341 AlaLeuSerThrThrArgGlyLeuIleProPheTyrGlyLysAlaIleProLeuGlnAlaIle 360
Db 2748 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTATC 2807
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCATCTTCTGTGATTAAGAAGAGAGCGCAGAACCTCGCGCA 2867
QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTGCATTTGGGCATCAATGCCGTGCTACTACCGCGCTTGAACGTGTCCGTC 2927
QY 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACAGCGGCGCATGTGTGCTGTGTGGCAACGATGCCCTCATGACCGGCTATCC 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGACCTGCACCTCGGATGATGACTGCATACGTGTGTGCCACCAAGCATGATTTCAGC 3047
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCTTACCTTTCACCATTTGAGACATCACCGCTCCCGAGATGCTGTCTCCGCACT 3107
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 3108 CAACGTGGGGGAGACTGAGAGGGGAGAACGAGCATTTACAGATTTGTGGACCCGGGG 3167
QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATATGTTGCACTGCTCGTCTGTGATGATGACGAGGCTGT 3227
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGTGATAGCTCACGCGCGGACGACATCAAGTTAGGTACGAGAGGTACATGACACC 3287
QY 521 ProGlyLeuAspProValCysGlnAspHisIleGluIlePheTrpGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCCGTGTGCGCAGAGACATTTGAATTTTGGAGGGCGCTTTTACAGGCTTC 3347
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysLeuLeuProTyr 560
Db 3348 ACTCATATATGATCCCATCTTTCATCCAGACAAGAGAGTGGGAGAACTTTCCTTAC 3407
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrPasp 580
Db 3408 CTGTAGCGTATCAACAGCACCGTGTGCGTACGAGGCTCAAGCCCTCCCATCATGTGGAC 3467

QY 581 GIMetTrpLysCysLeuIleAlaGluLeuYsProThrLeuHISGlyProThrProLeuLeu 600
DB 3468 CAGATGGGAAGGTTGATTGCTCAAGCCACCCCTCAATGGCCACACCCCTGCTA 3527
QY 601 TTAATGluGlyAlaValGlnAsnGluValThrLeuThiHisProValThiLysTyrIle 620
DB 3528 TACAGACTGGGCGCTGTTCAGATGAATACACCTGACGCAACCCATATCATTC 3587
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
DB 3588 ATGACATGATGTCGGCGGACCTGGAGAGTGTACAGACACCTGGGTGCTCTGGCGCG 3647
QY 641 ValLeuAlaIleAlaValAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 3648 GTCTGGGTGCTTTGGCGCGCTATTGCTGTCAACAGCTGCGGTGCTCATAGTGGCAGG 3707
QY 661 ILeValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 3708 GTGCTCTGTCTCGGGAAGCCGGCAATCATACCTGACAGGAAAGTCTCTACCGAGATTG 3767
QY 681 AspGluMetGluGluCys 686
DB 3768 GATGAGATGAAAGAGTGC 3785
RESULT 8
109331 7310 bp DNA linear PAT 02-DEC-1994
LOCUS Sequence 15 from Patent WO 8904669.
DEFINITION 109331
ACCESSION 109331 GI:587966
VERSION 109331.1 GI:587966
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 15 01-JUN-1989;
FEATURES
source 1..7310
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores: 2,54e-200 Length: 7310
Pred. No.: 3574.00 Matches: 672
Score: 99.42% Conservative: 10
Percent Similarity: 97.96% Mismatches: 4
Best Local Similarity: 98.78% Indels: 0
Query Match: 6 Gaps: 0
DB: US-09-930-591-2 (1-686) x 109331 (1-7310)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleIleThr 20
DB 1728 CTGGCGGCCATCAACGGGTACGCGCCAGACAAAGGCGCTCTCTAGGGTGCATATCAC 1787
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 1788 AGCTTAACCTGCGCGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCACTGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 1848 GCCCAACCTTCTCTGGGACAGTGCATCATGGGTGTGCTGACTGTCTTACCCAGGGGCC 1907
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 1908 GGAACGAGACCAATCGGTGTCACCAAGGTCTGTCTATCCAGATGATACCATGTAGAC 1967
QY 81 GlnAspLeuValGlyTTPProAlaProGlnGlyAlaAspSerLeuThrProCysTrpCys 100
DB 1968 CAAGACCTGTGGGTGGCGCTCCGACAGTATAGCGCTCATTTGACACCTGCACCTTGC 2027

QY 101 GlySerSerAspLeuTyrIleValThrArgHisAlaAspValIleProValArgArgArg 120
DB 2028 GGCTCTCCGACCTTATCTGTGTACAGACAGCCGATCTCATTCCTCCGCGCGCGCG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleLysGlySerSer 140
DB 2088 GGTGATAGCAGGGGCACTGCTGTGCGCCCGGCCCTTTCATCTTGAAGAGCTCTCTG 2147
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 2148 GGGGTCCGCTGTGTGCGCGCGGACGCGGTGGCATATTGAGCGCGCGGTGTC 2207
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 2208 ACCGTGATGCTGTAGAGCGGTGACCTTATCCCTGTGAGAACTTACAGAACACCATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 2268 AGTCCCGGTGTTCACGATATACCTCTCTCACAGTATGCCCCAGAGCTTCAGGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 2328 GCTCACTTCATGCTTCCACAGCAGCGGCAAAAGCAACAAGTCCGGCTGATATGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 2388 GCTCAAGGCTATAAAGTGTGATGACTCAACCCCTGTGTGCAACCTGGGCTTGTGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 2448 GCTTACATGTCCAAAGCTCATATGATGCATCTTACATCAAGAACCGGGGTGAAACAAT 2507
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 2508 ACCACCTGGACGCCCATCACTACCTCACTCACTGCGCAAGTCTCTTCCGACGCGGTGC 2567
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 2568 TCGGGGGGGCTTATACATATATTTGTGACAGTGCATCCACGGAATGCCACATCC 2627
QY 301 ILeLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 2628 ATCTTGGGATGGGCACTGCTTACCAAGACAGACGCGGGCGAGCTGTGTGTG 2687
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 2688 CTCGCCACCCGACCCCTCGGGCTCGTCACTGTGCCCCCATCCAACTCAGAGAGT 2747
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 2748 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAATGATTC 2807
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 2808 AAGGGGGGAGACATCTCATCTTCTGTCACTTCAAAAGAGAGGCAACGAACTCGCGCA 2867
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 2868 AAGCTGGTGGCATTCATCATGCGGTGCGGTCTTACACCGCGGTCTTGAAGTGTGCTG 2927
QY 401 ILeProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 2928 ATCCGACCAACGGCGAGTGTGTGTGTGTGCAACCATGCTCTCATGACCGGCTATACC 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 2988 GGCGACTTCGACTGGGTGATGACTGCATACGTTGTATCCAGACAGTGCATTTCCAGC 3047
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 3048 CTGACCTTACCTTACCATCATGAGACATACGCTCCCGCAGATGTCTGTCTCCGACT 3107

QY	61	GLYThrArgThrIleAlaSerProLysGlyProValIleGlnMetYrThrAsnValAsp	80
Db	1908	GGAAcGAGGACCAATCGCTCAcCGAAGGCTCTTCACTCAAGATTAACAAGTAAGAC	1967
QY	81	GLnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	1968	CAGACCTTGGGGCTGGCCCGCTCCGAAGGTAGCCGCTCATTTGACACCCGTGACCTTGC	2027
QY	101	GlySerSerAspLeuYrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
Db	2028	GGCTCCTCGACCTTTACCTGGTCAcGAGGACCGCGATGTCAATTCCTCGTGGCGCGG	2087
QY	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerYrLeuLysGlySerSer	140
Db	2088	GATGATAGCAGAGGGGAGCGCTGTGTGCGCCCGCCATTTCTACTTGAAAGGCTCTCG	2147
QY	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	2148	GGGGGTCCGCTGTTGTGTGCCCCGGGGGACCGCCGAGGACATATTAGAGGCGCGGTGTGC	2207
QY	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGlySerLeuGluThrThrMet	180
Db	2208	ACCCTGGAGTGGCTTAAGCGGTGACTTTATCCTGTGAGAACTTAAGAACCAATG	2267
QY	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal	200
Db	2268	AGGTCCCGGTGTTCAcCGATACCTCCTCAcCAGTAAGTGCACAGAGCTTCAAGGTG	2327
QY	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	2328	GCTCAcCTTCCAACTGCCCAcAGGCGGGCAAAACACCAAGTCCCGCTGCATATGCA	2387
QY	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	2388	GCTCAGGGCTAATAGGTGTACTCAACCCCTGTGTGTGCGAACACTGGGGTTTGGT	2447
QY	241	AlaYrMetSerLysAlaHisGlyLleAspProAsnIleArgThrGlyValArgThrIle	260
Db	2448	GCTTCAATGTCCAAAGGCTCATGGATCGATCCTTAACATCAGAGCCGGGGTGAACAATT	2507
QY	261	ThrThrArgLysProIleThrYrSerThrYrGlyLysPheLeuAlaAspGlyGlyCys	280
Db	2508	ACCACTGGGAGGCCCACTACGTACTTCACCTACGGCAAGTCTCTTCCGACGGCGGTGC	2567
QY	281	SerGlyGlyAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer	300
Db	2568	TCCGGGGGGCTTATATGACATATATTGTGTGACGATGCCACTCCACGAGATGCCACATCC	2627
QY	301	IleLeuGlyTlAcGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	2628	ATCTTGGGATGGGCACTGTCTCTTGAcCAGAAGACCTGCGGGGCGAGACTGGTTGTG	2687
QY	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	2688	CTCGGCAcCGGCAcCCCTCGGGGCTCCGTAcGTGAGCCCATGCCAATCGAGGAGGTT	2747
QY	341	AlaLeuSerThrThrGlyGluIleProPheYrGlyLysAlaIleProLeuGluAlaIle	360
Db	2748	GGCTGTGTCCACCAcCGGAGAGATCCCTTTTTCGCGAAGGCTATCCCTCGAAGTATTC	2807
QY	361	LysGlyGlyArgHisLeuIlePheCysHisSerIleLysValCysAspGluLeuAlaAla	380
Db	2808	AAAGGGGGAGACATCTCATCTTCTGTCTATTCAAGAAAGATGGGAGCAATCGCCGCA	2867
QY	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrYrArgGlyLeuAspValSerVal	400
Db	2868	AAAGCTGTGCAATTTGGGCACTCAATGCGTGGGCTACTACCGGAGCTTGAcGTGTCCGTG	2927
QY	401	IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr	420
Db	2928	ATCCCAcCAAGCGGCGCATGTGTGTCTGTGCGGCAcCGATGCCCTCATGACCGGCTAATCC	2987
QY	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440

Dd		2988	GCGCACTTCGACTGGATGAAGACTGCACAATACGGTGTGACCCGACAGCTCATTTGACC	3047
Oy		441	LeuAspProThrPheThrIleGIuThrIleThrlleuProGlnAspAlaValSerArgThr	460
Dd		3048	CTTAACCCCTACCTTCACCAATTGACAATAACGCTCCCGAGAATGCTGTCTCCGGACT	3107
Oy		461	GlnArgArgGlyArgTrnGlyArgGlyLysProGlyIleTyrrArgPheValAlaProGly	480
Dd		3108	CACGTGGGGGACAGACTGGCGAGGGGGAAAGCAGGAGCTCACAGATTGTGGACCCGGG	3167
Oy		481	GluArgProSerGlyMetPheAspSerValIleuCyseIuCystrYrAspAlaGlyCys	500
Dd		3168	GAGGCCCCCTCCGGCAATGTTGCACTCGTCCTCTGTGAGTCATGACGAGGCTGT	3227
Oy		501	AlaTrpTryGlyIleuThrProAlaGlnThrThrValArgLeuArgAlaTyrrMetAsnThr	520
Dd		3228	GCTTGATGAGCTCAGGCCCGGAGACTRCACTTAGCTTACGAGGCTACATGAACAC	3287
Oy		521	ProGlyIleuProValCyseGlnAspHisIleuGlnIuBheTrpGluGlyValPheThrGlyLeu	540
Dd		3288	CCGGGGCTCCCGTGTGCACAGACCATCTTGAAATTTGGGAGGGCGCTTTTACAGGGCTC	3347
Oy		541	ThrIstIleAspAlaHisPheIleuSerGlnThrTrpGlnSerGlyIuAsnIleuProTyrr	560
Dd		3348	ACTAATAAGATGCCCACTTTCTATCCGACACAAAGCAGAGTGGAGAACTTCCCTTAC	3407
Oy		561	IleuValAlaTyrrGlnAlaThrValCyseAlaArgAlaGlnAlaProProPheSerTrpAsp	580
Dd		3408	CTGGTAGGCTACCAAGCCACCGTGTGGCTAGGGCTCAAGCCCCTCCCGCATCTGGGAC	3467
Oy		581	GlnMetTrpIuCyseIleuIleArgIleuIuSerThrIleuHisGlyProThrProIleu	600
Dd		3468	CAGATGTGAAGTGTATTATTCGCTCAAGGCCACCTTCATGGAGCCAACCCCTGCTA	3527
Oy		601	TyrArgIleuGlyAlaValAlaGlnAspGluValThrIleuThrHisProValThrIysTrile	620
Dd		3528	THCAACTGGGGCGCTTCCAAATGMAATCACCTGACGACCCAGTACCAATATCATC	3587
Oy		621	MethTrpCysMetSerAlaAspIleuGluValAlaThrSerThrTrpValIleuValGlyGly	640
Dd		3588	ATGACATCATGTCGCGCCGACCTGAGGTCTGTCAAGACACCTGGTGGTCTGTGGCGGC	3647
Oy		641	ValIleuAlaAlaIleuAlaAlaTyrrCysIleuSerThrGlyCysValIleValIleValArg	660
Dd		3648	GTCCCTGGGCTGTGGCCCGCTATTGCTGTGCACAGGCTCTCGTGTCAATGTGGGACGG	3707
Oy		661	IleValIleuSerGlyLysProAlaIleIleProAspArgGluValIleuTyrrArgIuPhe	680
Dd		3708	GTGCTTGTGTCCGGGAAACCGGCATCATCACTGACAGGAAAGTCTCTTACCGAGATTTC	3767
Oy		681	AspGluMetGluGluCys	686
Dd		3768	GATGATGAGAAAGATGC	3785
	RESULT 10			
	LOCUS	108294	9185 bp	DNA linear PAT 02-DEC-1994
	DEFINITION	Sequence 1 from Patent EP 0388232.		
	ACCESSION	108294		
	VERSION	108294.1	GI:588994	
	KEYWORDS			
	SOURCE	Unknown.		
	ORGANISM	Unknown.		
	REFERENCE	Unclassified.		
	AUTHORS	1 (bases 1 to 9185)		
	TITLE	Houghton,M., Choo,Q.-L. and Kuo,G.		
	JOURNAL	NANBV diagnostics and vaccines		
	FEATURES	Patent: EP 0388232-A1 1 19-SEP-1990;		
		Location/Qualifiers		
		1..9185		
		/organism="unknown"		
		/mol_type="unassigned dna"		

ORIGIN

Alignment Scores:

Pred. No.: 3,366-200
Score: 3574.00
Percent Similarity: 99.42%
Best Local Similarity: 97.96%
Query Match: 98.78%

Length: 9185
Matches: 672
Conservative: 10
Mismatch: 4
Indels: 0
Gaps: 0

US-09-930-591-2 (1-686) x 108294 (1-9185)

QY 1 MetAlaProIleThraIaIaYrAlaGlnGlnThraArgIleuLeuGlyCysIleIeThr 20
DB CTGGCGCCCATTCACGGCGTACGCCACAGCAAGAGGGGCTCTAGGGTCATATACCC 3454
QY 21 SerLeuThrGlyArgAspIysAsnGlnValGluGlyGluValGlnIleValSerThra 40
DB AGCCTAACCTGCGCGGACAAACCAAGTGAAGGTACGATGTGTCAACTGCT 3514
QY 41 AlAGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThraIleYrHisGlyAla 60
DB GCCCAACCTTCTCTGGAGACGTGCATATGGGTGTGTGACTGTCTACAGGGGCC 3574
QY 61 GlyThrArgThrIleAlaSerProIyGlyProValIleGlnMetYrThrAsnValAsp 80
DB GGAACGAGACCAATCGCGTACACCAAGGTCCTGTCTATCCAGATGTATCAATGTAGAC 3634
QY 81 GluAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB CAAGACCTTGTGGGTGGCGCGCGCGCGCAAGTAGAGCGCTCATTTGACACCTTGCTGC 3694
QY 101 GlySerSerAspLeuYrLeuValThraArgHisAlaAspValIleProValArgArgArg 120
DB GGCCTCCCGACCTTTACCTGTGTACAGAGGACGCCCATGTCTATCCCGTGGCGGCGG 3754
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerYrLeuYrGlySerSer 140
DB GGTGATAGACAGGGGACCTGCTGTGCGCCCGGCTTCCTTCACTTGAAGGCTCTCG 3814
QY 141 GlyGlyProLeuLeuLeuProAlaGlyHisAlaValGlyIlePheAlaAlaValCys 160
DB GGGGGTGGGTGTGTGTGCGCGGGGACGCGTGGGCATATTAGGGCGCGGTGTGC 3874
QY 161 ThrArgGlyValAlaValAspPheIleProValGlnSerLeuGlnThrThrMet 180
DB ACCCGTGAAGTGGCTTACGGCGGTGACTTATTCCTTGTGAGAACCTTAGAACACCATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal 200
DB AGGTCCCGGTGTTCACGATACCTCTCTCCACCACTAGTGGCCCGACCTTCAGGTG 3994
QY 201 AlaHisIleuHisAlaProThrGlySerGlyLysSerThrIyValProAlaAlaIYrAla 220
DB GCTCACTCTCATGTCTCCACAGGACGGGCAAAAGCCAAAGTCCCGGCTGCATATGCA 4054
QY 221 AlaGlnGlyYrIyIysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB GCTCAGGGCTATAGGGCTAGTACTAACCCCTCTGTGTGTGCAACAGGGGCTTTGGT 4114
QY 241 AlaIyMetSerIyAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB GCTTACATGTCCAAAGGCTCATGGGATCATCTTACATCAAGACCGGGGTGAGAACATTT 4174
QY 261 ThrThrGlySerProIleThrYrSerThrYrGlyLysPheLeuAlaAspGlyGlyCys 280
DB ACCACGTGGACGCCCATACAGTACTCCACTACGGCAAGTTCTTGGCAGGGGGGTGC 4234
QY 281 SerGlyValAlaIyAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB TCGGGGGGGCGCTTATGATCATATATTGTAACAAGTCCACTCCAGAGTGCACATCC 4294
QY 301 IleuLeuGlyIleGlyThrValLeuAspGlnAlaGlyIuThraIaGlyAlaArgLeuThraVal 320

DB 4295 ATCTGGGCAATCGGCACTGCTTGTACCAAGACAGACTCGGGGGGACATGCTTGTG 4354
QY 321 LeuAlaThraIaThraProProGlySerValThraValProHisProAsnIleGluGluVal 340
DB CTGGCACCGCCACCCCTCGGGGCTCGTACGTGTGCCCCATCCCAATCGAGAGGATT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheYrGlyLysAlaIleProLeuGluAlaIle 360
DB GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGGAGGCTATCCCTCGAAGTATCC 4474
QY 361 IyGlyGlyYrArgHisIleuIlePheCysHisSerLysLysIyCysAspGlyIuLeuAlaIa 380
DB AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAGAGTCCGACGAATCCGCGCA 4534
QY 381 IyLeuValAlaIleuGlyValAsnAlaValAlaIleThraYrArgGlyLeuAspValSerVal 400
DB AAGCTGTGCAATTTGGCATCAATGCCGTGGCTTACTACCGCGTCTTGAAGTGTCCGT 4594
QY 401 IleProThrSerGlyAspValValAlaIaThraAspAlaLeuMetThrGlyPheThr 420
DB ATCCGACCAAGGGCGCATGTGTGTGTGTGGCAACGATGCCCTCATGACGGGTATACC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB GGGACCTTCACTCGGTGATGACGCAATACCTGTGTCAACCCAGACAGTGCATTTCCAGC 4714
QY 441 LeuAspProThrPheThrIleGlyIuThrIleThraLeuProGlnAspAlaValSerArgThr 460
DB CTTGACCTTACCTTCCACATTGAGACATACCTCCCGGAGATGTGTGTCCGCACT 4774
QY 461 GlnArgArgGlyYrArgThrGlyArgGlyLysProGlyIleYrArgPheValAlaProGly 480
DB CAACGTGCGGGGACGAGCTGGCAAGGGGGAAGCCAGGATCTACAGATTTGTGGACGGGG 4834
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysYrYrAspAlaGlyCys 500
DB GAGCGGCCCTCCGCGCATGTGTGCACTGCTGTCTGTGTAAGTATGACGAGGCTGT 4894
QY 501 AlaTrpTrpGluLeuThraProAlaGlyIuThrThraValArgLeuArgAlaYrMetAsnThr 520
DB GCTTGTATGAGCTTCACGCGCGCGGACACTACGTTAAGGTACGAGGTACATGAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisIleuGlnPheTrpGluGlyValPheThrGlyLeu 540
DB CCGGGGCTTCCCTGTGTGCGAGACCATCTTGAATTTTGGAGGGCGCTTTACAGGCTTC 5014
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProYr 560
DB ACTCATATATGATGCCACCTTCTATCCAGACAAACAGAGTGGGAGAACTTCTCTTAC 5074
QY 561 LeuValAlaIyTrpGlnAlaThraValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB CTGTATGCGTACCAACCAACCGGTGTGCGGTAGGGGTCAAGCCCTCCCTCATGTGGGAC 5134
QY 581 GluMetTrpLysCysLeuIleArgLeuLysProThrThraHisGlyProThrProLeuLeu 600
DB CAGATGTGAAGTGTGTTGATTCGCTTCAGGCCCACTCATGTGGCCACACCCCTGTCTA 5194
QY 601 YrArgLeuGlyValAlaValGlnAsnGluValThraLeuThraHisProValThryIyIle 620
DB TACAGACTGGGCGCTGTTCAGATGAATACACCTCAGACGACCCAGTCAACAAATACATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThraSerThrTrpValLeuValGlyGly 640
DB ATGACATGACATGTGTGGCCGACCTGAGGTGTGTACAGACACCTGGGTGTCTGTGGGCG 5314
QY 641 ValLeuAlaAlaLeuAlaAlaIyYrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB GTCTGTGCTCTTGTGGCGCGGTATTCCTGTCAACAGGCTGTGTATAGTGGGACG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuYrArgGluPhe 680

Db 5375 GTCGCTCTTGTCCGGGAAGCCGGCAATCACTGACAGGAGTCTTACCGAGATTTC 5434
QY 681 ASPGLUWEGUJUGUCYS 686
Db 5435 GATGAGATGGAAGAGTGC 5452

RESULT 11
BD091382 9185 bp DNA linear PAT 27-AUG-2002
LOCUS HCV cultivation method in eucaryotic cells.
DEFINITION BD091382
ACCESSION BD091382.1 GI:22636993
VERSION JP 2001314192-A/3.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 9185)
AUTHORS Weiner,A.J., Steimer,K.S. and Houghton,M.
TITLE HCV cultivation method in eucaryotic cells
JOURNAL Patent: JP 2001314192-A 3 13-NOV-2001;
CHIRON CORP
COMMENT OS Homo sapiens (human)
PN JP 2001314192-A/3
PD 13-NOV-2001
PF 15-MAR-2001 JP 2001075114
PR 25-AUG-1989 US 398667
PI AMY J WEINER, KATHELYN S STEIMER, MICHAEL HOUGHTON PC
CI 31N5/00, CI 2N5/10, CI 2N7/00//CI 2N7/00, CI 2N1/93, CI 2N15/00, PC
CI 2N5/00
CC HCV cultivation method in eucaryotic cells
FH Location/Qualifiers
FT source 1..9185
1..9185 /organism='Homo sapiens (human)'.
Location/Qualifiers
1..9185 /mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN

Alignment Scores:
Pred. No.: 3,366-200 Length: 9185
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
Gaps: 0
DB: 6

US-09-930-591-2 (1-686) x BD091382 (1-9185)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3395 CTGGGGCCCATCAACGGCGTACGCCCGACGACAAAGGGGCTCTTGGGGGATATCAAC 3454
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGlnValGlnIleValSerThrAla 40
Db 3455 AGCTTAACCTGGCGGAGCAAAACCAAGTGAAGGGTGAAGTCCAGATTGTGCACTGCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlnValCysThrThrValTyrHisGlyAla 60
Db 3515 GCCCAACCTTCTGCGCAACGTGATCAATGGGGTGTGGCTGTCTTACCAACGGGGCC 3574
QY 61 GlyThrArgThrIleLeuSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3575 GGAACAGAGACATCCCGTCAACCAAGGGTCCGTGATCAAGTATACCAATGTAGAC 3634
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlnValAlaArgSerLeuThrProCysThrCys 100
Db 3635 CAAGACCTTGTGGGGTGGCCCGCTCCGCAAGTGAAGCGCTCATGACACCTTGACATTGC 3694
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120

Db 3695 GGCCTCCGAGACTTTTACCTGTGCACAGGACCCGATGTCAATTCCTCCGCGCCGCGG 3754
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlySerSer 140
Db 3755 GGTGATACAGAGGGCAGCTCTGTGCGCCGCCCATTTCTTACTTGAAGGCTCTCCG 3814
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheAlaAlaValCys 160
Db 3815 GGGGGTCCGCTGTGTGCCCCCGGGGACCCCGTGGGCATATTAGGCGCGCGGTGTC 3874
QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3875 ACCCGTGAAGTGGCTTAAGCGGTGACCTTTATCCCTGCGAGAACCTAGAACACCATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3935 AGTCCCGGCTTTCACGGAATACCTCTCCACAGAGTGCCTCCCAAGCTTCAGGTG 3994
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3995 GCTCACTTCATGCTCCCAAGGACGGGCAAAAGCACCAAGTCCCGGCTGATATGCA 4054
QY 221 AlaGlnGlyTyrTyrValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4055 GCTCAGGCTTATTAAGTGTCTACTTCAACCCCTGTGTGCTGCAACACTGGGCTTGGT 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCCAAAGCTCATAGGATGATCGATCTTAACATCAGAGCGGGGTGAACAATT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACGTGGAGCCCACTACGATCTCCACTTACCGCAAGTTCTTGGCAGCGGGGTGC 4234
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGGGCTTATATGATCAATTAATTTGTGACGAGTGCACCTCCACGATGCCACATC 4294
QY 301 IleLeuGlyTyrLeuThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTGGGCAATCGGCACTGCTCTTGAACCAAGAGACCTGGGGGCGAGACTGTTGGT 4354
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTGCGCACCGCCACCTCCGGGCTCGGTACGTGCCCATCCAACTCGAGGGGT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4415 GCTGTGCCACCAACCGGAGAGATCCCTTTTACGGGAAGGCTATCCCTCGAAGTAAATC 4474
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLeuLysCysAspGlyLeuAlaIle 380
Db 4475 AAGGGGGGAGACATCTCACTTCTGTGATTTCAAGAAAGAGCGAGCAACTCGCGCG 4534
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4535 AAGCTGTGCGATTTGGCATCAATGCGGTGCTACTACCGCGGTCTTGAAGTGTGCTG 4594
QY 401 IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCGACCAAGCGGAGATGTGTGTGTGTGCAACGATGCCCTCATGACCGGCTATAC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GGCACCTTCGACTGGGTGATGATGCAATACGTTGTGACCCAGACAGTGCATTCACG 4714
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCTTACCTTACCATGATGAGACATACAGCTCCCAAGATGCTGTCTCCGCACT 4774
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyTyrIleTyrArgPheValAlaProGly 480
Db 4775 CAACGTGGGGGAGAGCTGGCAGGGGAGAGCGCAAGCATCTACATTTGTGGCAGCGGG 4834

QY 481 GUAArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTyrAspAlaGlyCys 500
DB 4835 GAAGGCCCCCTCCGCGAATGTTGCACTGCTCCGTCTGTAGAGCTATGACGACGGCTGT 4894
QY 501 AlaTyrTyrGlyLeuThrProAlaGlyThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4895 GCTTGGATATAGCTCAAGCCGCCGAGACTACAGTTAGCTACAGGTACATGAAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGlyValPheThrGlyLeu 540
DB 4955 CCGAGGCTTCCCGGTCGACAGACCACTTGAATTGGAGGGCGCTTTTACAGGGCTC 5014
QY 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 5015 ACTCATATATATGATGCCCTTCTATCCAGACAAAGAGAGTGGGGAGAACCTTCCCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
DB 5075 CTGGTAGCTGACCAAGCAGCCGCTGCTAGGGCTCAAGCCCTCCCATCGTGGGAC 5134
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGTGAGAGTGTGATTGGCTTCAAGCCCACTTCATGGGCCCAACACCCCTGCTA 5194
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 5195 TACAGACTGGCGCTGTTCAGATGAAATCATCCTCGACGACCCAGTCAACAAATATATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTyrValLeuValGlyGly 640
DB 5255 ATGACATGATGATGTGCGCGACCTGAGAGTGTCTACAGACACCTGGGCTCGTGGGGCG 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrLysCysValValIleValGlyArg 660
DB 5315 GTCTGTGCTCTTGGCGCGGTATGTGCTGTCAACAGCTGCGGTGATCATAGTGGAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 5375 GTCTGTCTTGTCCGGGAGCCGGCAATCATCTGACAGGAGAGTCTTACCAAGAGTTC 5434
QY 681 AspGluMetGluGlyCys 686
DB 5435 GATGAGATGGAAGAGTGC 5452

RESULT 12
ARI66930 LOCUS 9379 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6284249.
ACCESSION ARI66930
VERSION ARI66930.1 GI:16243325
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9379)
AUTHORS Barban, V.
TITLE Fusion polypeptide having the C protein and B1 protein of hepatitis C virus
JOURNAL Patent: US 6284249-A 1 04-SEP-2001;
FEATURES Location/Qualifiers
source 1..9379
/organism="unknown"
/mol_type="unassigned DNA"

ALIGNMENT Scores:
Pred. No.: 3,45E-200 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x ARI66930 (1-9379)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3395 CTGGCGCCATTCAGCGGGTACGCCGACAGACAGAGGGGCTCTTAGGGTCAATATACACC 3454
QY 21 SerLeuThrGlyYArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 3455 AGCTTAACCTAGCGCGGAGCAAAACCAAGTGAAGGAGGTCCAGATTGTGTAACGTCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
DB 3515 GCCCAAACTTCTCTGCAACGTGCATCATAGGGGTGTGCGACTGTCTTACACAGGGGCC 3574
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3575 GGAACAGAGACATCCGCTACCCAAAGGCTCTGTATCCAGATGATACAAATGTAGAC 3634
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3635 CAAAGCTTGTGGGCTGGCGCGCTCCGCAAGTAGCGGCTCATTTACACCTGCACATTGC 3694
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 3695 GGCTCTCGGACCTTACCTGATCAGAGGACAGCCGATGTCAATCCCGTCCGCGCGCG 3754
QY 121 GlyAspGlyYArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 3755 GGATATAGCAAGGGGAGCGCTGCTGCGCCCGCCCAATTTCTTCACTTGAAAGGCTCTCG 3814
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyYIlePheArgAlaAlaValCys 160
DB 3815 GGGGGTCCGCTGTGTGTGCCCCGGGGCAAGCCGTGGGCAATTTAGGGCGGGGTGTGC 3874
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3875 ACCCGTAGAGTGTGCTAAGGGGTGAGCTTATCTCTGGAGAGAACCTAGACAAACCATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3935 AGGTCCCGGTGTTCACAGATATCTCTTCCACAGATATGCCCCAGACTTCCAGAGT 3994
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3995 GCTCACTCTCATCTCTCCACAGGCAAGCGGAAAGACCAAGTCCCGGCTGCATATGCA 4054
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4055 GTCAGAGGCTATAGGCTAGTACATCAACCTCTGTGTCTGCAACATCGGGCTTTGGT 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 4115 GCTTACATGTCCAAAGGCTCATGGGATCGATCTTACATCAAGACCGGGGTGAACAACATT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 4175 ACCACTGGCAGCCCATCATCACTCTCACCTTACAGGCAAGTCTTCCGACCGCGGGTGC 4234
QY 281 SerGlyValAlaTyrAspIleIleIleLysAspGluCysHisSerThrAspAlaThrSer 300
DB 4235 TCGGGGGGGCTTATGACATATATTTGGAGAGTGCACATCCACGAGTGCACATCC 4294
QY 301 IleLeuGlyYIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValArgLeuThrVal 320
DB 4295 ATCTTGGGATCGGCACTGTCTTGAACAGAGACATCGGGGGCGAGACAGTGTGTG 4354
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4355 CTGGCCACCGCACCCCTCGGGCTCGTCACTGTGCCCATCCAAATCGAGAGGTT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360

Accession	Gene	Protein	Length	Start	Stop	Strand	ORF	Start	Stop	Strand	ORF
AB01300	gltC	GltC	4415	1	4415	+	100%	1	4415	+	100%
AB01300	lysG	LysG	361	1	361	+	100%	1	361	+	100%
AB01300	magS	MagS	4475	1	4475	+	100%	1	4475	+	100%
AB01300	lysE	LysE	381	1	381	+	100%	1	381	+	100%
AB01300	amcT	AmcT	4535	1	4535	+	100%	1	4535	+	100%
AB01300	ilp	Ilp	401	1	401	+	100%	1	401	+	100%
AB01300	atfC	AtfC	4595	1	4595	+	100%	1	4595	+	100%
AB01300	glaP	GlaP	421	1	421	+	100%	1	421	+	100%
AB01300	ggaC	GgaC	4655	1	4655	+	100%	1	4655	+	100%
AB01300	leuA	LeuA	441	1	441	+	100%	1	441	+	100%
AB01300	glnA	GlnA	461	1	461	+	100%	1	461	+	100%
AB01300	caaG	Caag	4775	1	4775	+	100%	1	4775	+	100%
AB01300	glnA	GlnA	481	1	481	+	100%	1	481	+	100%
AB01300	gagG	GagG	4835	1	4835	+	100%	1	4835	+	100%
AB01300	alaT	AlaT	501	1	501	+	100%	1	501	+	100%
AB01300	gctG	GctG	4895	1	4895	+	100%	1	4895	+	100%
AB01300	proG	ProG	521	1	521	+	100%	1	521	+	100%
AB01300	ccgG	CcgG	4955	1	4955	+	100%	1	4955	+	100%
AB01300	thrI	ThrI	541	1	541	+	100%	1	541	+	100%
AB01300	actC	ActC	5015	1	5015	+	100%	1	5015	+	100%
AB01300	leuA	LeuA	561	1	561	+	100%	1	561	+	100%
AB01300	ctgG	CtgG	5075	1	5075	+	100%	1	5075	+	100%
AB01300	glnM	GlnM	581	1	581	+	100%	1	581	+	100%
AB01300	cgaG	CgaG	5135	1	5135	+	100%	1	5135	+	100%
AB01300	tyrA	TyrA	601	1	601	+	100%	1	601	+	100%
AB01300	tyrA	TyrA	5195	1	5195	+	100%	1	5195	+	100%
AB01300	metH	MetH	621	1	621	+	100%	1	621	+	100%
AB01300	atgA	AtgA	5255	1	5255	+	100%	1	5255	+	100%
AB01300	valE	ValE	641	1	641	+	100%	1	641	+	100%
AB01300	gltC	GltC	5315	1	5315	+	100%	1	5315	+	100%
AB01300	lysE	LysE	661	1	661	+	100%	1	661	+	100%
AB01300	gltC	GltC	5375	1	5375	+	100%	1	5375	+	100%
AB01300	aspG	AspG	681	1	681	+	100%	1	681	+	100%
AB01300	gatG	GatG	5435	1	5435	+	100%	1	5435	+	100%

DEFINITION	SEQUENCE 1 FROM PATENT US 6538123.
ACCESSION	AR301300
VERSION	AR301300.1
KEYWORDS	GI:31689076
ORGANISM	Unknown.
SOURCE	Unknown.
REFERENCE	1 (bases 1 to 9379)
AUTHORS	Baraban, V.
TITLE	Vaccine composition for preventing or treating hepatitis C
JOURNAL	Patent: US 6538123-A 1 25-MAR-2003;
FEATURES	Location/Qualifiers
source	1..9379
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
Alignment Scores:	
Pred. No.:	3,45e-200
Score:	3574.00
Percent Similarity:	99.42%
Best Local Similarity:	97.96%
Query Match:	98.78%
DB:	6
	Gaps: 0
US-09-930-591-2 (1-686) x AR301300 (1-9379)	
QY	1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB	3395 CTrGGGCCCATCATCAGCGCGTACCCCGACGACGACGAGGGGCGCTTACGGGTGATATCACC 3454
QY	21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB	3455 AGCTTAACTCGCGCGGACAAACCAAGTGAAGGTGAGTCCAGATGTGTCAACTGCT 3514
QY	41 AlaGlnThrPheIleuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB	3515 GCCCAAACTTCTGGCAACGTGACATCATGGAGGTGTCTGGACGTGTACACGGGGCC 3574
QY	61 GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB	3575 GGAAACAGAGCAACATCCCGTACCCCAAGGGTCTCTGTCATCCAGATGATATACAAATGATAC 3634
QY	81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB	3635 CAAAGACTTGTGGGTGGCGCCGCTCGCAAGGAGCGGCTCATGACACCTGCACCTTGC 3694
QY	101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB	3695 GGCTCTCGGACCTTAACTGGTGCAGAGGACACGCGAATGCTATTCGCCGTGGCGGGCG 3754
QY	121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB	3755 GGTGATACGAGGGGACACCTGCTGTGGCCCGGCCCATTTCTCATTGAAGGCTCTCG 3814
QY	141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheThrGlnAlaValCys 160
DB	3815 GGGGGTCCGCTGTGTGCCCGCGGGGACACGCGTGGCATATTAGAGCGCGGGGTGTC 3874
QY	161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGlnThrThrMet 180
DB	3875 ACCCGTGAAGGTGAAGCGGTGAGCTTTATCCCTGTGAGAACTAGAGCAACATCAG 3934
QY	181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB	3935 AGGTCCCGGCTTCAAGGATATCTCTCCACAGATAGTGCCTCCAGAGCTTCCAGGTG 3994
QY	201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrTyrValProAlaIleTyrAla 220
DB	3995 GCTCACTCCATGCTCCACAGGAGGCAAAAGCCAAAGTCCCGGCTGATATGCA 4054
QY	221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIleThrMetGlyPheGly 240

Db 4055 GCTCAGGGCTAATGAGTGTAGTCAACCCCTGTGGTCTGCACACTGGGCTTTGGT 4114
 Qy 241 AATATMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4115 GCTTAATGTGTCAGAGCTCATGGATGATCTTACATCAGACCGGGGTGAACAACTT 4174
 Qy 261 ThrThrGlySerProIleThrIleThrIleThrIleThrIleThrIleThrIle 280
 Db 4175 ACCACTGGCAGCCCATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 4234
 Qy 281 SerGlyValAlaIleAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4235 TCGGGGGGCGCTTATGACATTAATTTGTGACAGAGCCATCTCCAGAGTGCACATTC 4294
 Qy 301 IleGluGlyIleGlyThrValIleAspGlnAlaGluThrAlaGlyAlaGlyLeuThrVal 320
 Db 4295 ATCTTGGGCATCGGCACTGCTTGCACAGAGAGCTCGGGGGGGAACCTGTTGGT 4354
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 4355 CTGCGCACCGCCACCCCTCGGGGCTCGTCACTGTGCCCATCCCAATCGAGAGGTT 4414
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyValAlaIleProLeuGluAlaIle 360
 Db 4415 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGGCAGGCTATCCCTCGAAGTATTC 4474
 Qy 361 LysGlyGlyValArgHisLeuIlePheCysHisSerIleValIleValIleValIleVal 380
 Db 4475 AAGGGGGGAGACATCTCATCTCTCATCTCATCTCATCTCATCTCATCTCATCTCATCT 4534
 Qy 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaIleThrIleValIleAspValSerVal 400
 Db 4535 AACCTGGTGCATTTGGGCATCAATGCCGTGGCTACTACCGCGTCTTGAAGTGTCCGTC 4594
 Qy 401 IleProThrSerGlyAspValValValValAlaAlaThrAspAlaLeuMetThrIlePheThr 420
 Db 4595 ATCCCGACCGAGCGGCGATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4654
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 4655 GGGGACTTCGACTCGGATGAGTCACTGCAATACGTGTCTCAACCCAGACAGTGAATTCAGC 4714
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4715 CTGACCCCTACCTTCACTCACTGAGACATCACTGCTCCCGAGATGCTGTCTCCGCACT 4774
 Qy 461 GlnArgArgGlyArgThrGlyValGlyValIleValIleValIleValIleValIleVal 480
 Db 4775 CAACGTGCGGCGAGACTGCGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4834
 Qy 481 GluArgProSerGlyMetPheAspSerValLeuGluCysGlyArgAspAlaGlyCys 500
 Db 4835 GAGGCGCCCTCCGCGCATGTTGACTGCTCGCTCTGTAGAGTCAATGACGAGGCTGT 4894
 Qy 501 AlaTrpIleGluLeuThrProAlaGluThrThrValArgLeuArgAlaIleValIleVal 520
 Db 4895 GCTTGGATGAGCTCAAGCCCGCGAGACATCACTTACGCTACGAGGCTACGATGAACACC 4954
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValIlePheThrGlyLeu 540
 Db 4955 CCGGGGCTTCGCGGTGCGCAGACCATCTGAATTTGGGAGGCGCTTTTACAGGCGCTC 5014
 Qy 541 ThrHisIleLeuAspAlaHisPheLeuSerGlnThrIleGlnSerGlyValIleLeuPheThr 560
 Db 5015 ACTCAATTAAGATGCCCTTTCTATCCAGCAAAAGCAAGTGGGAGAACCTTCTTAC 5074
 Qy 561 LeuValAlaIleValIleValIleValIleValIleValIleValIleValIleValIleVal 580
 Db 5075 CTGTTACGTCATCAAGCAACCGTGTGCGTACGAGGCTCAAGCCCTCCCATCGTGGGAC 5134
 Qy 581 GlnMetTrpIysCysLeuIleValGluLeuIleValProThrIleHisGlyIleProThrProLeu 600
 Db 5135 CAGATGTGAAGTGTGTGATTCGCTCAAGCCACCTCCATGGGCGCAACCCCTGTCTA 5194

Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIleValIle 620
 Db 5195 TACAGACTGGGCGCTTTCAGATGAATATCAACCTGACCAACCACTCACTCACTCACT 5254
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrThrValIleValIleVal 640
 Db 5255 ATGACATGATGATGCGGCGCACTCGAGGTCGTACAGACCACTGGGTGCTCGTTGGCGGC 5314
 Qy 641 ValLeuAlaAlaLeuAlaIleValIleValIleValIleValIleValIleValIleVal 660
 Db 5315 GTCTTGCTGCTTGGCGCGCGATTCCTGTCAACAGGCTGCTGCTGCTGCTGCTGCTGCT 5374
 Qy 661 IleValLeuSerGlyIleAspProAlaIleIleIleProAspArgGluValLeuThrArgGluPhe 680
 Db 5375 GTCTGCTTGTCCGGGAAAGCGGCAATCATCTGACAGGGAAGTCTCTTACGAGATTC 5434
 Qy 681 AspGluMetGluGluCys 686
 Db 5435 GATGAGATGAGAGATGTC 5452

RESULT 14
 AR176483 9401 bp DNA linear PART 17-DEC-2001
 LOCUS AR176483
 DEFINITION Sequence 9 from patent US 6312889.
 ACCESSION AR176483
 VERSION AR176483.1 GI:17918838
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES
 1 (bases 1 to 9401)
 Houghton, M., Choo, Q.-L., and Kuo, G.
 Combinations of hepatitis C virus (HCV) antigens for use in
 immunosays for anti-HCV antibodies
 Patent: US 6312889-A 9 06-NOV-2001.
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..9401
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 3,466-200 Length: 9401
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: Gaps: 0

US-09-930-591-2 (1-686) x AR176483 (1-9401)
 Qy 1 MetAlaProIleThrAlaIleGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 3417 CTGGCGCCCATCAACCGCGTACGCCAGCAAGAGGCGCTCTTACGGTGCATTAATCACC 3476
 Qy 21 SerLeuThrGlyArgAspIleAsnGlnValGluGlyValGlnIleValIleValIleVal 40
 Db 3477 AGCTTAATCTGGCGGAGCAAAACCAAGTGAAGGAGTCCAGATTGTGTCAACTGCT 3536
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleValIleVal 60
 Db 3537 GCCCAACCTTCTTGGCAACGTCATCAATGGGTTGCTGAGCTGTCTACACGGGGCC 3596
 Qy 61 GlyThrArgThrIleAlaSerProIleValIleValIleValIleValIleValIleVal 80
 Db 3597 GGAACGAGGACATCCGTCACCAAGGCTCTGTATCCAGATGATACCAATGTAGAC 3656
 Qy 81 GluAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 3657 CAGACCTTGTGGGCTGGCCGCTCCGCAAGTAGACCGCTCATTTGACACCTTCGACTTC 3716
 Qy 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArgArg 120

```

Db      3717 GGCCTCTCGGACCTTTACCTGTCACGAGCAGCGAGTGCATTCCTCCGTGGCCGGCGG 3776
Qy      121  G1YAspG1YArgG1YSerLeuLeuSerProAlaArgProIleSerGlyLeuGlySerSer 140
Db      3777 GGTGATGCGAGGGGCGAGCCCTGCTGCGCCCGCCCAATTCCTACTTGAAGAGCTCTCG 3836
Qy      141  G1G1YProLeuLeuSerProAlaG1YH1a1aValaG1Y1ePheH1a1a1aValaCys 160
Db      3837 GGGGGTCCGCTGTTGTGTCCCGGGGGCAAGCCGTGGGCATATTATAGGCGCGGCTGTGC 3896
Qy      161  ThrArgG1YAla1a1aValaValaAspPhe1eProValG1YSerLeuG1YH1a1a1a 180
Qy      3897 ACCCGTGAAGTGGCTAAGCGGTGAGCTTATCCCTGTGGAACCTAGAGACAAACATG 3956
Db      181  ArgSerProValaPheSerAspAsnSerSerProAlaValaProG1YSerG1YArgVala 200
Db      3957 AGGTCCCGGCTTTCACGAGTAACCTCCTCCACAGTACTGACCCCAAGCTTCAAGTG 4016
Qy      201  AlaH1a1a1a1a1aProH1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 220
Db      4017 GCTACACCTCCATGCTCCCAAGCAGCGGCAAAAGCAAGGTCGCCGCTGATATGCA 4076
Qy      221  AlaG1YArg1YArg1YValaValaLeuValaLeuAsnProSerVala1a1a1a1a1a1a 240
Db      4077 GCTCAGGGCTATTAAGGTGCTAGTACTCAACCCCTGCTGTGCAACAACCTGGCTTGT 4136
Qy      241  AlaTyrMetSerLeuValaH1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 260
Db      4137 GCTTACATGTCACAGGCTCATGGAGTGCATCTTAACATCAAGACCGGGGTGAGAACAT 4196
Qy      261  ThrThrG1YSerProIleThrTyrSerThrTyrG1YArgPheLeuAlaAspG1YG1Y 280
Db      4197 ACCATGCGACGCGCCATCAGTACGTAACCTCAAGCAAGTCTTGCAGACGCGGGTGC 4256
Qy      281  SerG1YG1YAlaTyrAsp1e1e1e1eCysAspG1YCysH1a1a1a1a1a1a1a1a1a 300
Db      4257 TCGGGGGGCGCTTATGACATTAATATTGTGACAGTGCACCTCAGCGAGTGCACATCC 4316
Qy      301  IleLeuG1Y1eG1YThrValaLeuAspG1YAlaG1YH1a1a1a1a1a1a1a1a1a 320
Db      4317 ACTTGGGATCGGACGCTGCTTACCAAGCAGGACCTGGGGGCGAGCTGTTGTG 4376
Qy      321  LeuAla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 340
Db      4377 CTGCGACCGCCACCCCTCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4436
Qy      341  AlaLeuSerThrThrG1YG1YLeuLeuProPheTyrG1YArg1YAla1eProLeuG1Y 360
Db      4437 GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAGATATC 4496
Qy      361  LysG1YG1YArgH1a1a1a1eLeuLeuPheCysH1a1a1a1a1a1a1a1a1a1a1a 380
Db      4497 AAGGGGGGAGACATCTCATCTTCTGTCTTAAGAAGAAGTGGACAGAACTCCGCCA 4556
Qy      381  LysLeuVala1a1a1eG1YValaAsnAlaVala1a1a1a1a1a1a1a1a1a1a1a 400
Db      4557 AACCTGTGTCATTTGGGATCATGCGCGGCTTACTACCGGAGCTTGAACGTGCCGTC 4616
Qy      401  IleProThrSerG1YAspValaValaVala1a1a1a1a1a1a1a1a1a1a1a1a 420
Db      4617 ATCCGACCGAGCGGATGTTGTCTGTGCGCAACGAGGACCTCATGACCGGCTATACC 4676
Qy      421  G1YAspPheAspSerVala1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 440
Db      4677 GGGGACTTCGACGCGGATGAGCTGATACATGCTGTCTCAACCAAGCAATTCATTCAGC 4736
Qy      441  LeuAspProThrPheThr1eG1YH1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 460
Db      4737 CTTGACCTTACCTTCAACATTAAGACATACAGCTCCCGAGGATCTGTCTCCGACCT 4796
Qy      461  G1H1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 480

```

```

Db      4797 CAACGTGCGGGCAGAGCTGGCAGGGGAGAGCCAGCATCTTACAGATTGTGTGACCGGGG 4856
Qy      481  G1H1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 500
Db      4857 GAGGCGCCCTCCGCGAGTTCGATCGTCCGCTCTGTGATGCTATGACGAGGCTGT 4916
Qy      501  AlaTyrG1YLeuLeuThrProAlaG1YH1a1a1a1a1a1a1a1a1a1a1a1a1a 520
Db      4917 GCTTGTATGAGCTCAAGCCCGCGAGACTAGCTAGGCTACAGAGGCTACATGAAAC 4976
Qy      521  ProG1YLeuProValaCysG1YAspH1a1a1a1a1a1a1a1a1a1a1a1a1a 540
Db      4977 CCGGGGCTTCCCGTGTGCAAGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCTC 5036
Qy      541  ThrH1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 560
Db      5037 ACTCATATAGATGCCCACTTCTTATCCAGCAAGAGAGTGGAGAACTTCTTAC 5096
Qy      561  LeuVala1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 580
Db      5097 CTGTAGGCTACCAAGCAGCCGTGTGCGCTAGGGCTCAAGCCCTCCCATCGTGGAC 5156
Qy      581  G1H1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 600
Db      5157 CAGATGTGAGAGTTTATATGCTCAAGCCCAACCTCCATGGGCAACCCCTGCTA 5216
Qy      601  TyrArgLeuG1YAla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 620
Db      5217 TACAGATCGGGCGCTTCAAGATGAATACACCTGAGCGCAAGCCCAAGTCAACAAATAC 5276
Qy      621  MetThrCysMetSerAlaAspLeuG1YVala1a1a1a1a1a1a1a1a1a1a1a 640
Db      5277 ATGACATGATGATCGGCGACCTGAGAGTGTGACAGACAACTGGTGTCTGCTGGCGG 5336
Qy      641  ValLeuAla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 660
Db      5337 GTCTGTGCTGCTTGGCCCGCGATTCGCTGCAACAGGCTGCTGATATGTTGGGCGG 5396
Qy      661  IleValLeuSerG1YAspProAla1a1e1eProAspArg1YValaLeuTyrArg1Y 680
Db      5397 GTGCTGTGTCCGGAGAGCGGCAATCATACCTGACAGGAGAGTCTTACGAGAGTTC 5456
Qy      681  AspG1YMetG1YLeuG1YCys 686
Db      5457 GATGAGATGAGAGAGTGC 5474

RESULT 15
LOCUS      E66593              9401 bp    DNA          linear    PAT 18-UN-2001
DEFINITION Hepatitis C virus asialoglycoprotein.
ACCESSION   E66593
VERSION     E66593.1 GI:13026028
KEYWORDS    JP 1999071395-A/1.
SOURCE      Hepatitis C virus
ORGANISM    Hepatitis C virus
VIRUSES:    ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirus.
REFERENCE   1 (bases 1 to 9401)
AUTHORS    Robert,O.R., Frank,M., Kent,B.S., Barbara,A.J. and John,A.H.
TITLE      Hepatitis C virus asialoglycoprotein
PATENT     JP 1999071395-A 1 16-MAR-1999;
JOURNAL     CHIRON CORP
OS          Hepatitis C virus
PN          JP 1999071395-A/1
PD          16-MAR-1999
PF          14-APR-1998 JP 1998103178
PR          08-NOV-1990 US          611,965, 09-NOV-1990 US          611,419 PR
PI          13-SEP-1991 US          758,880
PI          ROBERT O BARUSUTON, FRANK MARCUS, KENT B SUDIAMU, PI BARBARA A
JABAST, JOHN A HALL
PC          C07K14/18, A61K39/29, G01N33/576//C12N15/09, C12P21/02, (C12P21/02, PC
C12R1:91),

```


Db	5277	ATGACATGCATGTCGGCCGACCTGGAGGTCTGTCACGACACCTGGGTCTGCTGTCGCGCC	5336
QY	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5337	GTCTGTGCTGCTTGGCCGGATATGCTCTGTCAACAGGCTGGGTGTCAATAGTGGCAGG	5396
QY	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5397	GTCTCTTGTCCGGGAAGCCGCATCATTACTGACAGGGAAGTCTTACCGAAGTTTC	5456
QY	681	AspGluMetGluGluCys	686
Db	5457	GATGAGATGGAAGATGC	5474

Search completed: September 17, 2004, 10:30:07
 Job time : 7286 secs

PI Salberg M;
XX WPI, 2002-339446/37.
DR P-PSDB; AAE21837.
XX Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
PS Claim 1; Page 64-65; 90pp; English.
XX
CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is a DNA encoding HCV NS3/4A protein
XX
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,44e-245 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-930-591-2 (1-686) x AAD34500 (1-2061)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysAlaIleThr 20
DB 1 ATGGCGCTATCAAGCGCTATGCCAGACAGAGGGCCCTTTGGGATGATATATACCC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGlnGlyGluValGlnIleValSerThrAla 40
DB 61 AGCTTGAACCGGCGCGGACAAACACAGGTGAGGGGTGAGTTGATCTGTCAACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysAlaIleAsnGlyValCysThrThrValTyrHisGlyAla 60
DB 121 GCCCAGACTTCTTGGCACTGCACTTACAGGGGGGTGGAGTGTCTGCTTACCATGGAGCC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 181 GGAACAGAGACCATTTGGCTGCTACCTAAGGCTCTGTATTCAGAGTACACCAATGTGGAC 240
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlnValAspSerLeuThrProCysThrCys 100
DB 241 CAAGCCTCTGTAAGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACCAATGACATTCG 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGCTCTCTGACCTTTTACCTGTGTACAGAGGACGCCCATGTCAATTCCTGTGCGCCGACG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 361 GGTATGCGCAGGGGAGGAGCTGCTTTCGCCCGGCTATCTCTACTTAAAGGCTCCCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 421 GGAGGCCCTCTGCTGTGTGCCCGCAGACATGCCGTAGGCAATATTCAGGCGCGGTATGC 480
QY 161 ThrArgGlyValAlaAlaValAlaAspPheIleProValGlnSerLeuGluThrThrMet 180
DB 481 ACCCTGGAGTGGCTTAAGGCGGTGACTTCACTCCCGTAGAGGCTTGAAGACCAACATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTCCCGGTGTCTTCAGACAACTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 601 GCCCACTGACGTCTCCACCGGACGCGTAAAGACCAACAGGTCCGCGCGCATACGCA 660

QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGCTACAGAGTGTGGTGTCTCAACCCCTCGCTTGCAACAATGGGCTTTGGT 720
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCCAAAGGCCCATGGATGATTCCTTAACATCAGACCTGGGGTGAAGCAAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 781 ACTACTGGAGCCCGATCAAGTATTCACCTACCGGAAGTTCTTGCACCGCGGGTGT 840
QY 281 SerGlyValAlaTyrAspIleIleLeuCyAspGluCysHisSerThrAspAlaThrSer 300
DB 841 TCAGGGGGGTCTTATACATATATATTTGTGACGAGTGCACATCCAGATGACACATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValArgLeuThrVal 320
DB 901 ATCTTGGCATTTGGCACTGTCTCTTGAACAGCAGACCGCGGGCGAGACTGACTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGluVal 340
DB 961 CTGCGACCGCTACCCCTCGGGCTCCGTCACGTGTGCCCATCTTAACATCAGAGAGTT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1021 GCTCTGCCACTACCGGAGAGATCCCTTTATGACAAAGCTATTCCTTGAAAGCAAT 1080
QY 361 LysGlyValArgHisLeuIlePheCysHisSerLysLysLysCysAspGlyLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCATCTTCTGCACCTCAAAAGAGAGTGCACAGGCTGGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1141 AAACGTGTGCGTGGCGGTCAATGCCGTGCTTACTACCGGGCTTATGTGTCTCGTC 1200
QY 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCGACAGATGGTGCCTTGTGCTGTGGCAACAGCGCCCTCATGACCGGCTTTTAC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 1261 GGGGACCTTCATTCGGTGAATGAGCTGCAACAGTGTGTACCCAGACATGCACTTCACG 1320
QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCTTACCTTACCATTTGAGACATACAGCTTCCAGAGATGTGTCTCCGTA 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTGGGGTGAAGCTGGCAGAGGGAGCCAGGCAATCAAGATTGTGGACCGGGG 1440
QY 481 GlnAspProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1441 GAGGCTCTTCTGGCAATGTTGACTGTCTGTCTCTGCAAGTGTCTATGACGGGGTGT 1500
QY 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGTATGATGCTTACGCCCGCGAGACACAGTATAGCTTACATGAAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
DB 1561 CCGGAGCTTCCCGGTGCCAAAGCAATCTTGAATTTTGGAGGGCGCTTTTACGGGCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1621 ACCCAATTAACGCGCCCTTCTATCCCAACAAGAGAGTGGGAAACCTTCCCTAT 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
DB 1681 CTGTATACGTAACCAAGCAACGTGTGTGCTTAAGAGCTCAACCCCTCCCGTGTGGAC 1740

QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 1741 CAGATGTGGAAGTCTGATCCGTCTCAGCCCACTTCATGGGCCAACAACCTTCTGCTA 1800
 QY 601 TyrArgLeuGlyAlaValGlnAenGluValThrLeuThrHisProValThrLysTyrIle 620
 DB 1801 TATAGACTGGGCGCTGTCAGAAATGAAATGACCTTACGACCACTGATCAACAGTATATC 1860
 QY 621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTrpValLeuValGlyGly 640
 DB 1861 ATGACAGTATGTCGTGACCTGAGAGTGTGTCACGAGTACCTGGGTCTGTTGGGGCC 1920
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 DB 1921 GTTCTGGCTCTTGGCGCGGATTTGCTATCCACAGGCTGGGTGTCAATAGTAGTAGG 1980
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 DB 1981 ATTGTCTTGTCCGAAAGCCGCAATCATACCGACAGGAAGTCTCTACCGGAGATTC 2040
 QY 681 AspGluMetGluGluCys 686
 DB 2041 GATGAATGGAAGAGTGC 2058
 RESULT 2
 AAD31767 standard; DNA; 2061 BP.
 ID AAD31767
 XX AAD31767;
 AC 18-JUN-2002. (first entry)
 DT
 XX
 DE Hepatitis C virus (HCV) NS3/4A DNA coding region.
 XX
 DE Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
 KW cytolethal; immunostimulant; vaccine; ribavirin; immune response; cancer;
 KW de.
 XX
 OS Hepatitis C virus.
 XX
 FT Key Location/Qualifiers
 FT 1..2061
 FT CDS /*tag= a
 /product= "HCV NS3/4A protein"
 XX
 PN MO200213855-A2.
 XX
 PD 21-FEB-2002.
 PD
 PF 15-AUG-2001; 2001MO-IB001808.
 PF
 PR 17-AUG-2000; 2000US-0225767P.
 PR 29-AUG-2000; 2000US-0229175P.
 PR 03-NOV-2000; 2000US-00705547.
 PR
 PA (TRIP-) TRIP AB.
 PA
 XX
 PI Salberg M, Hultgren C;
 XX
 DR WPI; 2002-241837/29.
 DR P-PSDB; AAE19900.
 DR
 XX
 PT Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 PT
 XX
 PS Claim 1; Page 94-95; 120PD; English.
 XX
 CC The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to a
 CC hepatitis C antigen in humans, domestic, sport or pet species and as

CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is HCV NS3/4A DNA coding region
 XX
 SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,44e-245 Length: 2061
 Score: 3618.00 Matches: 686
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x AAD31767 (1-2061)
 QY 1 MetAlaProIlePheThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1 ATGGCGCTATTCAGCGGCTATGCTCCAGACAGAGGAGGCTTTGGAGTGCATATAC 60
 QY 21 SerLeuThrGlyArgAspLysAenGlnValGluGlyValGlnIleValSerThrAla 40
 DB 61 AGCTTACCGGCGCGGACAAACACAGTGGAGGTGAGTTCAATCGTCACTGCT 120
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAenGlyValCysTrpThrValTyrHisGlyAla 60
 DB 121 GCCCAGACTTCTTGGCAACCTGCAATTAAGGGGTGTGGAGCTGTACATGAGAGCC 180
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAenValAsp 80
 DB 181 GGAACAGAGCACATTCGTCACCTAAGGATCTGTATCCAGATGACCAATGTGGAC 240
 QY 81 GluAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 241 CAAAGACTGTAGGCTGGGCGGCTCCCAAGGCGCCGCTCATTAACACATGACTTGC 300
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 DB 301 GGCTCTCGGACCTTACCTGATGTCAGAGGACGAGATGTCATTCCTGTGGCCGACGG 360
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProAlaArgProIleSerTyrLeuLysGlySerSer 140
 DB 361 GGTGATGGCAGGGGCGGCTGCTTGGCCCGGCTCATCTTAAAGGCTCTCTCG 420
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 421 GGAAGCCCTCTGCTGTGCCCCCGACGACATGCCGTAGGCAATTCAGAGCCGGATGC 480
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGlnThrMet 180
 DB 481 ACCGATGAGTGGCTAAGGGGTGAGCTTCATCCCGTAGAGGCTTAGAGAACCATG 540
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 541 AGGTCCCGGTGTCTCAGACATCTCTCCCAACAGAGTGGCCCAAGGCTCAAGTG 600
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 601 GCCCAGCTCATGCTCCACCGGACGGAGTAAAGAGCAACAGGTCCCGGCGGATAGCGA 660
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 661 GCTCAAGGCTTACAAAGTGTGCTGCTCAACCTCTGCTGTGCTCAACATGGCTTTGGT 720
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAlaIleArgThrGlyValArgThrIle 260
 DB 721 GCTTACATGTCCAAAGCCCATGAGTATGATCTTAACATCAGGAGCTGGGTGAGGACAAAT 780
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 781 ACTACTGGAGCCCGATTCAGTATTCACCTTCAGGAGAGTCTTCCACAGCGGAGGT 840
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300

```

Db      841 TCAGGGGGTGTATGACATATATATTGTGACAGAGCCACTCCAGGATGCAACATCC
Qy      301 ILeuenclytleqlythrrvalleuaspglnalagluhrrlaaglyalaaglyethrrval
Db      901 ATCTTGGGCATTTGGACATGTCCTTGACCAAGACACCGGGGGGAGACATGACATGTCG
Qy      321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal
Db      961 CTGCCACCGCTACCTCCGGGCTCCGTCATCTGTGCCATCTTAACATCGAGGAGGTT
Qy      341 AlaLeuSerThrThrGlyGluIleProPheTyGlyLysAlaIleProLeuGluAlaIle
Db      1021 GCTCTGTCCACTACCGAGAGATCCCTTTATGGCAAGGCTATTCCTTGAAGCAATT
Qy      361 LysGlyGlyArgHisLeuIlePheCysHisSerIleValLysCysAspGluLeuAlaAla
Db      1081 AAGGGGGGAGACATTCATCTCTGCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      381 LysLeuValAlaLeuGlyValAsnAlaValAlaValAlaThrArgGlyLeuAspValSerVal
Db      1141 AAATGCTGCGGTGGGCGGTCAATGCGGTGCTTACACCGCGGCTTGATGTGTCGTC
Qy      401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr
Db      1201 ATCCCGACCAAGTGTGACGTGTGTCGTGCACTGACCGCTCATGACCGGCTTTAAC
Qy      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGluThrValAspPheSer
Db      1261 GGCACATTCGATTCGGATGAGACTGCAACAGTGTGTCACCGACAGATGTCATTCAC
Qy      441 LeuAspProThrPheThrIleGluThrIleThrLeuProGluAspAlaValSerArgThr
Db      1321 CTGACCTTACCTTACCATGAGACATACATCGCTTCCCGAGATGCTGCTCCGACT
Qy      461 GluArgArgGlyArgThrGlyArgGlyLysProGlyIleTyraGpPheValAlaProGly
Db      1381 CAACGTGGGGGTAGAGACTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyraAspAlaGlyCys
Db      1441 GAGGCTCTTTCGACATGTTGACTGCTGCTCTCTGCGAGTCTATGACCGGGTGT
Qy      501 AlaTrpTyrgluLeuThrProAlaGluThrThrValArgLeuArgAlaTyraMetAsnThr
Db      1501 GCTTGGATATGAGCTTACCGCCGAGAGACACAGTTAGGCTTACAGAGCATATGAAACC
Qy      521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu
Db      1561 CCGGAGACTTCCCGTGTGCGCAAGACATCTGAAATTTGGAGGGCGCTTTTACCGGTC
Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyx
Db      1621 ACCCAATAGACGCCCACTTCTTATCCAGACAAAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      561 LeuValAlaTyrglnAlaThrValCysAlaArgAlaGlnAlaProProSerTyraAsp
Db      1681 CTGGATAGGATACCAAGCACCGTGTGCGCTAGACTCAAGCCCTCCCGCTCGTGGAGC
Qy      581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu
Db      1741 CAGATGTGGAAGTGTGATCGTCTCAAGCCCACTCCATGAGGAGAGAGAGAGAGAGAG
Qy      601 TyraArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrlle
Db      1801 TATGACTGGGCGCTGTCCAAATGAATGATCACTCGAGCACCAGTCAACCAATATATC
Qy      621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyraValLeuValGlyGly
Db      1861 ATGACATATATGTGGCTGAGCTGAGAGGTGCTCAAGAGTACCTGGGTGCTGTTGGCGGC
Qy      641 ValLeuAlaAlaLeuAlaAlaTyraCysLeuSerThrGlyCysValValIleValGlyArg

```

```

Db      1921 GTTCTGCTGCTTGTGGCCGGATTCCTATCCACAGGCTGCGTATAGTAGTAGG
Qy      661 ILeValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyraArgLuphe
Db      1981 ATTGTCTTGTCCGGAAGCCCGCAATCATACCGAGAGAGAGAGAGAGAGAGAGAG
Qy      681 AspGluMetGluGluCys 686
Db      2041 GATGAATGGAAGAGTGC 2058

RESULT 3
AAB0868 standard; DNA; 2061 BP.
ID AAB0868;
AC AAB0868;
XX
XX 15-JAN-2004 (first entry)
XX
XX Hepatitis C virus NS3/4A DNA.
DE
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide; de.
XX
XX Hepatitis C virus.
OS
XX Key Location/Qualifiers
XX FT 1. .2061
XX FT CDS /*tag= a
XX FT /product= "Hepatitis C virus protein"
XX
XX US2002136740-A1.
XX
XX 26-SEP-2002.
XX
XX 15-AUG-2001; 2001US-00923955.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX
XX (SALT/) SALTBERG M.
XX PA (HULT/) HULTGREN C.
XX
XX SalIberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX DR P-PSDB; ABW00351.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX Claim 1; Page 60-61; 0pp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
XX antigen, where the antigen is derived from a hepatitis C virus. The vaccine
XX is useful in enhancing the immune response to a hepatitis C antigen where
XX the composition is delivered to an animal identified as requiring an
XX enhanced immune response. The vaccine is useful in the treatment and
XX prevention of hepatitis C infection. The present sequence is Hepatitis C
XX virus NS3/4A DNA
XX
XX Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2,44e-245 Length: 2061
XX Score: 3618.00 Matches: 686
XX Percent Similarity: 100.00% Conserves: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 9 Gaps: 0
XX
XX US-09-930-591-2 (1-686) x AAB0868 (1-2061)

```

Qy 1 MetAlaProIleThraIaIaIaGlnGlnThraGlyLeuLeuGlyCysIleIleThr 20
Db 1 ATGGCGCTATACAGGCTATAGCCACAGACAAGGGGCTTTGGAGTCAATATACCC 60
Qy 21 SerLeuThrGlyArgAspIysAsnGlnValGluGlyValGlnIleValIleSerThrAla 40
Db 61 AGCTTACCGGCGGAGCAAAACACAGGTGAGGTTCAGATCGTGTCAACTGCT 120
Qy 41 AlagIlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIYrHisGlyAla 60
Db 121 GCCACGCTTTCTTGGACCTGATTAACGGGTGTGTTGACTGTCTACATGAGACC 180
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetYrThrAsnValAsp 80
Db 181 GGAACAAGAGACATGGGTGCTACCTTAAGGGTCTGTATCCAGATGTACCAATGTGAC 240
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGTGCCTCATTTAACACATGACACTTGC 300
Qy 101 GlySerSerAspLeuYrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 301 GGCTCTCTGCACTTTACCTGTGTCACAGGACGCCATGTCACTTCTGTGCGCCAGCG 360
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerYrLeuLysGlySerSer 140
Db 361 GGTAATGGCAGGGGCACTGCTTTCGCGGCTATCTTACTTGAAGGCTCTCG 420
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGAGGCTCTGCTGTGCTCCCGCAGACATGCGTAGGCATATTCAAGCGCGGTATGC 480
Qy 161 ThrArgGlyValAlaIysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 481 ACCGTGAGTGGCTTAAGCGGTGACTTATCCCTGAGAGACTTAAGACACACATG 540
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal 200
Db 541 AGGTCCCGGCTTCTCAGACAACTCTCCACACAGACATGCGCCAGACTACCAAGTG 600
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaIYrAla 220
Db 601 GCCCACTGCAATGCTCCACCGGCAAGGATGAAGACCAAGATCCGCGCGCATAGCA 660
Qy 221 AlagIlnGlyYrLysValLeuValLeuAsnProSerAlaAlaIleThrMetGlyPheGly 240
Db 661 GCTCAGGAGCTACAGGTGCTGCTCAACCCCTCGTGTGCAACAATGGGCTTGGT 720
Qy 241 AlaYrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCCAAAGGCCATGGGATTGATCTTAACATCAAGACTGGGGGTGAGACAT 780
Qy 261 ThrThrGlySerProIleThrYrSerThrYrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 781 ACTACTGCGAGCCGATCAGTATTCACCTACCGCAAGTCTCTTGCAGCGGGGTGT 840
Qy 281 SerGlyGlyAlaYrAspIleIleIleCysAspGlyLysHisSerThrAspAlaThrSer 300
Db 841 TCGAGGGGTGCTTATGACATTAATTTGTGACAGTGCACCTCAGAGATGCAACATCC 900
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlyLysThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTTGGGCAATTGGCACTGTCTTGAACAAGACAGACCGGGGGGAGACTGACTGTG 960
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 961 CTCGCGACCGCTACCTCCGCGCTCCGCTCACTGTGCCCATCTTAACATCGAGAGAGTT 1020
Qy 341 AlaLeuSerThrThrGlyGluIleProPheYrGlyLysAlaIleProLeuGluAlaIle 360
Db 1021 GCTCTGTGCTACCTACCGAGAGATCCCTTTATGCAAGGCTATTTCCCTTGAAGCAATT 1080
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaIa 380

Db 1081 AAGGGGGAGAGATTCATCTTGTGCACTCAAAAGAGGACAGACTGCGCGCA 1140
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaYrYrArgGlyLeuAspValSerVal 400
Db 1141 AACTGTGCGCTTGGCGCTCAATGCGGTGCTTACTACCGGCTTATGTGTCCGT 1200
Qy 401 IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCGACCAAGTGTGACGTGTGCTGTGAGCAATGACGCGCTCATGACCGGCTTACC 1260
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1261 GGGACTTGATCGGTGTGATGACTGCACACGTGTGTACCCAGACAGTGCATTACG 1320
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCTTACTTTCACATTTAGACAAATCACGCTTCCCGAGATCTGTCTCCGTACT 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleYrArgPheValAlaProGly 480
Db 1381 CAACGTGCGGGTAGACTGGCAGAGGAAAGCAGCATCAAGATTGTGGCAGCGGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGlyCysYrYrAspAlaGlyCys 500
Db 1441 GACGCTCTTCTGGCATGTTTGACTGTGTCTGTGAGAGTATACCGGGGTGTG 1500
Qy 501 AlaTrpYrGluLeuThrProAlaGluThrThrValArgLeuArgAlaYrMetAsnThr 520
Db 1501 GCTTGGTAGAGCTTACGCCCGCCAGACACAGTTAGGCTACGACATACATGAACACC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 1561 CCGGACTTCCCGTGTGCCAAGACATCTTGAATTTTGGAGGGCGCTTTACGGGCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysLeuProYr 560
Db 1621 ACCACATGACCCGCCACTTCTATCCAGACAAACAGGTGGGAAAACTTCCCTAT 1680
Qy 561 LeuValAlaYrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 1681 CTGTGTGCGTACCAACCCACCGTGTGCTAGAGCTCAAGCCCTTCCCGTGTGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGAAGTGTGATCCGTCTCAAGCCACCTCTCATGGGCAACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrLysYrIle 620
Db 1801 TATAGACTGGGCGCTGTCCAGATGAAGTCAACCTGACGACACCAAGTATATC 1860
Qy 621 MetThrCysMetSerLysAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGATGTGCGCTGACCTGAGGTGTGTACCAAGTACCTGGGTGTGCTGTGGCGG 1920
Qy 641 ValLeuAlaAlaLeuAlaAlaYrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGCTGCTTGTGGCCGCTGATTCCTATCCACAGGCTGCTGATAGTAAAGTAA 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuYrArgGluPhe 680
Db 1981 ATTGTCTTGTCCGAAAGCCGCGCAATCATACCCGACAGGAGTCTCTACCGGAGTTC 2040
Qy 681 AspGluMetGluGluLys 686
Db 2041 GATGAATGGAAGATGC 2058

RESULT 4
AAN92106
ID AAN92106 standard; DNA; 7310 BP.
XX
AC AAN92106;
XX

DT 25-MAR-2003 (revised)
 DT 02-MAR-1990 (first entry)
 XX Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
 DE clones K9-1 through 15c.
 XX
 KM Hepatitis C virus; HCV, non-A, non-B hepatitis; NANBH.
 XX
 OS Hepatitis C virus.
 XX
 XX Key Location/Qualifiers
 FH CDS 3..7310
 FT /*tag= a
 XX
 XX EPJ18216-A.
 XX
 XX 31-MAY-1989.
 PD
 XX 18-NOV-1988; 88EP-00310922.
 PF
 XX 18-NOV-1987; 87US-00122714.
 ER 30-DEC-1987; 87US-00139886.
 PR 26-FEB-1988; 88US-00161072.
 PR 06-MAY-1988; 88US-00191263.
 PR 26-OCT-1988; 88US-00263584.
 PR 14-NOV-1988; 88US-00271450.
 XX
 XX (CHIR) CHIRON CORP.
 PA (CHIR) CHIRON CORP.
 XX
 XX Houghton M, Choo QL, Kuo G;
 DR WPI, 1989-159274/22.
 DR P-PSDB; AAP92050.
 XX
 XX Purified hepatitis C virus - and associated nucleic acids and
 PT polypeptide(s).
 XX
 XX Claim 3; Fig 47-1 - 47-8; 139pp; English.
 PS
 XX It is a double-stranded nucleotide sequence of the open reading frame
 CC (ORF) (tag a) extending through clones K9-1 to 15c of hepatitis C virus
 CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
 CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)
 CC it encodes could be used as immunoassay reagents and vaccines and to
 CC generate antibodies useful in diagnosis and passive immunotherapy for HCV
 CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC
 SQ Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 1,42e-241 Length: 7310
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 1 Gaps: 0
 US-09-930-591-2 (1-686) x AAN92106 (1-7310)
 QY 1 MetalaproteinThraAlaTyraAlaGlnGlnThraArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1728 CTGGGGCCCATCATCGGCGTACGCCAGCAGACAAAGGGGCTCTAGGGTGCATATCATCC 1787
 QY 21 SerLeuThraGlyArgAspIysAsnGlnValGlnGlyValGlnIleValSerThrAla 40
 DB 1788 AACCTAACTGGCGGAGCAAAAACCAAGTGAAGGAGTCAAGATTGTGTCATCTGCT 1847
 QY 41 AlaGlnThraPheLeuAlaThrCysIleAsnGlyValCysTrrThraValTyriSglVala 60
 DB 1848 GCCCAAACTTCTCTGGCAAGTGCATCATGGGGTGTGTGTGAGACTGTCTACACAGGGGCC 1907

QY 61 GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyThrAsnValAsp 80
 DB 1908 GGAACAGAGACCATCGCGTACCCCAAGGGTCCGTTCATCCAGATGATATACCAATGTAGAC 1967
 QY 81 GlnAspLeuValGlyTrrProAlaProGlnGlyAlaAspSerLeuThraProCysThrCys 100
 DB 1968 CAAGACCTTGTGGTGGCCGCTCCGAGAGTACCGCTCATTTACACCTGCACTTGC 2027
 QY 101 GlySerSerAspLeuTyrlleuValThraGlnIleAlaAspValIleProValArgArg 120
 DB 2028 GGCTCTCGGACCTTACCTGTCACGAGGACCGCGATGTCATTCGGTGGCGCGCGG 2087
 QY 121 GlyAspGlyArgGlySerleuLeuSerProAlaGProIleSerTyrlleuGlySerSer 140
 DB 2088 GGTGATAGCAGGGGAGCGCTGCTGTCCCGGCCCATTTCTTACTTGAAGGCTTCTCG 2147
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 2148 GGGGGTCCGCTGTGTGCCCCCGGGGACGCGCGTGGCATTTTAGGGCCGGGTGTGC 2207
 QY 161 ThrArgGlyValAlaIysAlaValAspPheIleProValGlySerLeuGlnThrThrMet 180
 DB 2208 ACCCGTGAAGTGGCTAAGCGGAGCTTATCCCTGTGAGAACCTAGAGCAACATG 2267
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrlleuVal 200
 DB 2268 AGGTCCCGGTGTGTACGATTAATCTCTCCACACGATGTCGCCAGAGCTTCACGATG 2327
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyIysSerThraValProAlaAlaTyraAla 220
 DB 2328 GCTCACTTCATGCTCCACAGGACGGGCMAAAGCAAGGTCGCCGCTCATATGCA 2387
 QY 221 AlaGlnGlyTyrlleuValleuValleuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 2388 GCTCAGGCTATGAAGGTGCTAGTACCAACCCCTGTGTGTCGCAACACTGGGCTTTGGT 2447
 QY 241 AlaTyrMetSerIysAlaHisGlyIleAspProAsnIleArgThraGlyValArgThrIle 260
 DB 2448 GCTTACATGTCAGAGCTCATGGATGATCTTACATGACGACCGGGTGAACAAT 2507
 QY 261 ThrThrGlySerProIleThrTyrlleuThraGlyIysPheLeuAlaAspGlyIysCys 280
 DB 2508 ACACAGTGGACGCCCATCATCGTACTCCACTACGGCAAGTTCCTTCCACGCGGGTGC 2567
 QY 281 SerGlyValAlaIysAspIleIleIleCysAspGlnCysHisSerThrAspAlaThrSer 300
 DB 2568 TCCGGGGGCGCTTATGACATTAATATTGTGACAGATGCCACTCCAGGATGCCAATCC 2627
 QY 301 IleLeuGlyTlleGlyThrValleuAspGlnAlaGluThraAlaGlyIleArgLeuThrVal 320
 DB 2628 ATCTGGGCAATCGGCACTGCTTTCACCAAGACGAGACTCGGGGCGAGACTGTGTG 2687
 QY 321 LeuAlaThraAlaThrProProGlySerValThraValProHisProAsnIleGlnGluVal 340
 DB 2688 CTGGCAACGCCCAACCCCTCGGGCTCGTCACTGTGCCCATCCCAACATCGAGAGGT 2747
 QY 341 AlaLeuSerThrThrGlyGlnIleProPheTyrlleuValAlaIleProLeuGlnAlaIle 360
 DB 2748 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGCAAGGACTATCCCTCGAAGTATC 2807
 QY 361 LysGlyIysIysIleuIlePheCysHisSerIysIysCysAspGlnLeuAlaIle 380
 DB 2808 AAGGGGGGAGACATCTCATCTTCTGTCAATTAAGAGAGAGCGCAAGATCCGCGCA 2867
 QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrlleuArgGlyLeuAspValSerVal 400
 DB 2868 AAGCTGTGCGCATGGCATCAATGCCGTGGCTTACTACCGCGTGTGACGCTGTCCGTC 2927
 QY 401 IleProThrSerGlyAspProValAlaValAlaThraAspAlaLeuMetThrGlyPheThr 420
 DB 2928 ATCCGACCAAGCGCGATGTGTGTGTGTGCAACGATGCCCTCATGACCGGCTATACC 2987
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440


```
Db 2988 GGGGACTTGCAGCTCGGATAGACTGCAATACGTGTCTACCCAGACAGTCGATTTCCAGC 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCTTACCTTACCACTTGAACAATACGCTCCCTCCAGATGCTGCTCCCGACT 3107
Qy 461 GlnArgArgIleArgThrArgIleArgIleArgProGlyIleArgPheValAlaProGly 480
Db 3108 CAACGTGGGGGAGGAGCTGGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3167
Qy 481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyGlyArgIleArgAlaGlyCys 500
Db 3168 GAGGCGCCCTCCGCGCATGTCGACTCGCTCCCTGAGGCTATGACGCGGCGCTGT 3227
Qy 501 AlaTrpTrpGluLeuThrProAlaGluThrValArgLeuArgAlaTyrMetAspThr 520
Db 3228 GCTTGATTAAGCTCAGCCCGCCGAGACATACAGTTAGGCTACGAGCGTACATGACACC 3287
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCCGTGTCCAGAACCATCTTGAATTTGGAGGGCGCTTTACAGGCTC 3347
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyIleAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCCACTTCTATCCAGCAAGACAGAGTGGGAGAACTTCTTAC 3407
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProPheTrpAsp 580
Db 3408 CTGGTACGTAACCAAGCACCAGTGTGGCTAGGGCTCAAGCCCTCCCACTGTGGAC 3467
Qy 581 GlnMetTrpIleCysLeuIleArgLeuIleProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGAAGTCTTGTGATTCGCTCAAGCCCACTTCAAGGCGCAACCCCTGCTA 3527
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValIleTyrIle 620
Db 3528 TACGAGCTGGGCGCGTGTTCAGAAATGAATACACCTGACGACACCAATATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValIleValGly 640
Db 3588 ATGACATGCAATGTGGCGGAGCTGGAGGTGTCTCAAGACCTGGGGGCTGTGGGGG 3647
Qy 641 ValLeuAlaIleAlaValAlaTyrCysLeuSerThrGlyCysValIleValGlyArg 660
Db 3648 GTCTGCTGCTTGTGGCGCGTATTCCTGTCAACAGGCTGCTGTCAATAGTGGGAGG 3707
Qy 661 IleValLeuSerGlyIleProAlaIleIleProAspArgIleValLeuTyrArgGluPhe 680
Db 3708 GTGCTTGTTCGGGAGGCGGCAATCATACCTGACAGGGAAGTCTTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGlyCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785
RESULT 5
ID AAN90336 standard; DNA, 7310 BP.
XX
AC AAN90336;
XX
DT 25-MAR-2003 (revised)
DT 19-JUL-2001 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Composite hepatitis C virus (HCV) cDNA.
XX
KM Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.
XX
OS Pan troglodytes.
XX
PN GB2212511-A.
XX
```

```
PD 26-JUL-1989.
XX
PF 18-NOV-1988; 88GB-00027024.
XX
PR 18-NOV-1987; 87US-00122714.
XX
PR 30-DEC-1987; 87US-00139886.
XX
PR 26-FEB-1988; 88US-00161072.
XX
PR 26-OCT-1988; 88US-00263584.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-215054/30.
XX
DR P-PSDB; AAP90288.
XX
PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
PS Disclosure; Fig 47; 30pp; English.
XX
CC The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by
CC aligning clones K9-1 through 15e in 5'-3' direction. The cDNA encodes
CC antigens which react with antibodies in patients with non-A non-B
CC hepatitis (NANBH). The cDNA can be used to design probes, or to
CC synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to
CC raise antibodies for immunosassay or treatment, or to produce vaccines.
CC See also AAP90288, and AAN90303-35. (N.B. This record was resubmitted to
CC correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR
CC field.)
XX
SQ Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No. 1,42e-241 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
US-09-930-591-2 (1-686) x AAN90336 (1-7310)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 :::: 1728 CTGGCGCCCATTCACGCGCGTACGCCAGACAAAGGGGCTCTTACGGTGCATTAATCACC 1787
Qy 21 SerLeuThrGlyArgAspIleAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1788 AGCTTAATGCGCGGGAACAAACCAAGTGAAGGTGAAGTCCAGATTGTGTCACTGCT 1847
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleArgIleAla 60
Db 1848 GCCCAAACTTCTGCGCAAGTGCATCAATGGGGTGTGCTGCGACTGTCTACACAGGGGCC 1907
Qy 61 GlyThrArgThrIleAlaSerProIleArgProValIleGlnMetIleThrAsnValAsp 80
Db 1908 GGAACGAGACCATTCGCTACCAAGGCTCGTATCCAGATGATACCAATGTAGAC 1967
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAAGACCTTGTGGGCTGGCCCGCTCCGCAAGTACGCTCATATGACACCTGCACTTGC 2027
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAlaArgArg 120
Db 2028 GGCTCTCGGACCTTACCTGATCAGAGGACGAGATGTATTCCTCCGCGCGCGGCGG 2087
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProAlaGlyProIleSerTyrLeuIleGlySerSer 140
Db 2088 GGTGATAGCAGGGGCGAGCTGTGCTGCGCCCGCCCAATTTCTATGAAAGGCTCTCTGC 2147
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
```

```

Db      2148 GGGGGTCCGCTGTTGTGCCCCCGGGGAGCGCGTGCGCATATTTAAGGGCGCGGTGTGC 2207
Qy      161  ThrArgIyValAlaIyValAspPheIleProValGluSerLeuGluThrThMet 180
Db      2208 ACCGCTGAGAGGCTTAAGCGGTGACTTTATCCCTGTGAGAACATAGAACACATG 2267
Qy      181  ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGluVal 200
Db      2268 AGGTCCCGGTGTTTCAAGGATTAATCTCTTCCACCAAGTAGTCCCGAGCTTCCAGGTG 2327
Qy      201  AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db      2328 GCTCACCTCCATGTCTCCACAGGAGCGGCAAAAGACCAAGGTCCTCCGCTGATATGCA 2387
Qy      221  AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db      2388 GCTAGGGCTATTAAGGTGCTAAGTCAACCCCTCTGTGTGCAACATGAGGCTTGTGT 2447
Qy      241  AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db      2448 GCTTACATGTCCAAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 2507
Qy      261  ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db      2508 ACCACGTGCGAGCCCATCACTACTCCACCTACGGAAGTTCTTGGCCGACGCGGGTGC 2567
Qy      281  SerGlyGlyAlaTyrAspIleIleIleGlyAspGluCysHisSerThrAspAlaThrSer 300
Db      2568 TCGGGGGGCGCTTATGACATATATTTGTGACAGAGTCCACTCAGAGATGCCACATCC 2627
Qy      301  IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db      2628 ATCTTGGGCACTGGCACTGTCTTGACCAACAGAGACTGGGGGGGAGACTGTGTGTG 2687
Qy      321  LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      2688 CTCGCCACCGCCACCCCTCCGGGCTCCGTCCTGCTGCGCCCATCCCAATCGAGAGAGTT 2747
Qy      341  AlaIleuSerThrThrGluIleGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      2748 GCTCTGTCCACCGAGAGATCCCTTTTACGCAAGGCTATCCCTCGAAGTATTC 2807
Qy      361  LysGlyGlyArgHisIleLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db      2808 AAGGGGGGAGACATCTCATCTTCTGTCTATCAAAAGAAAGATCGACGAATCGCCGCA 2867
Qy      381  LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      2868 AAGCTGTGCGCATTTGGGCATCAATGCCGTGCTTACTACCGCGGTCTTGACGTGTCCGTC 2927
Qy      401  IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      2928 ATCCCGACACCGCGGATGTGTGCTGCTGGCAACCAATGCCCTCATGACCGGCTATATCC 2987
Qy      421  GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      2988 GCGGACTTCGACTCGGTGATGACATCGCAATACGTGTGCAACCGACAGCTGATTTGACG 3047
Qy      441  LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      3048 CTTGACCTTACCTTCAACATTCAGACATACGCTCCCGCCAGATGCTGTCTCCGCACT 3107
Qy      461  GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db      3108 CAAGCTCGGGGAGAGACTGCGAGGGGAGGAGCAAGCATTCACATTTTGTGGCACCGGGG 3167
Qy      481  GlnArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTyrAspAlaGlyCys 500
Db      3168 GAGGCGCCCTCCGGGATGTTGACTCGTCCGCTCTGTGATGCTATGACGAGCGCTGT 3227
Qy      501  AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuAlaGlyAlaTyrMetAsnThr 520

```

```

Db      3228 GCTTGTGATGAGCTACGCCCCCGGAGACTACAGTTAAGCTACAGCGATCATGAACCC 3287
Qy      521  ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluIleValPheThrGlyLeu 540
Db      3288 CCGGGGCTTCCGCTGTGCGCAGAGACATCTTGAAATTTTGGAGGGCGCTTTTACAGGCTC 3347
Qy      541  ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db      3348 ACTCATATATATGCCCATCTTATCTTATCCAGACAAACAGAGTGGGAGAACTTCTCTTAC 3407
Qy      561  LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProPheSerPheAsp 580
Db      3408 CTGTATGCGTACCAAGCCACCGGTGTGCTAGGGCTTCAAGCCCTTCCCATCGTGGAC 3467
Qy      581  GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      3468 CAAGTGTGAAAGTGTGATTCGCTTCAAGCCCACTCCATGGGCCAACCCCTGTCTA 3527
Qy      601  TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      3528 TACAGACTGGCGCTGTTCAGATGAATACACCTGACGACACCATGCAACCAATATCATC 3587
Qy      621  MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db      3588 ATGACATGATGTGTGCGCGCACTGAGAGTGTACAGAGACCTGTGGTGTCTGTGGCGGC 3647
Qy      641  ValLeuAlaIleLeuAlaIleAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      3648 GTCTGCTGCTGCTTGTGGCCCGGTATTCCTGTCTCAACAGGCTGCTGTATGAGGGCAGG 3707
Qy      661  IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db      3708 GTCGCTTGTTCGCGGAAGCCGGAATCATCTGACAGAGGAGAACTCTTACCGAGAGTTC 3767
Qy      681  AspGluMetGluGluCys 686
Db      3768 GATGAGATGGAAGAGTGC 3785

RESULT 6
AA098221
ID  AA098221 standard; cDNA to mRNA; 7310 BP.
XX
AC  AA098221;
XX
DT  25-MAR-2003 (revised)
DT  15-AUG-1996 (first entry)
XX
DB  Hepatitis C virus clone genome.
XX
KM  Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;
    antibodies; immunoprophylaxis; sera; serum; ds.
OS  Hepatitis C virus.
XX
EN  US5443965-A.
XX
PD  22-AUG-1995.
XX
PF  05-APR-1991; 91US-00681703.
XX
PR  06-APR-1990; 90US-00505611.
PR  09-OCT-1990; 90US-00594854.
XX
PA  (GENE-) GENELABS INC.
XX
PI  Moeckli R, Reyes GR, Kim JP;
XX
XX  WPI; 1995-302120/39.
XX
DR  New nucleic acids encoding hepatitis C virus antigens - used to develop
PT  prods. for detection of HCV-infected sera and prodn. of vaccines and anti
PT  -HCV antibodies.
XX

```

PS Example 4, Fig 11, 71pp; English.

XX Hepatitis C virus (HCV) antigens can be used for detecting HCV infected
CC sera and individuals infected with HCV. They can also be used in an anti-
CC HCV vaccine or for the production of anti-HCV antibodies which can be
CC used for passive immunoprophylaxis. The antigens consistently identify
CC more HCV positive serum samples with a high degree of specificity. See
CC AA098202-14 and AA81939-51. (Updated on 25-MAR-2003 to correct PR
CC field.) (Updated on 25-MAR-2003 to correct PR field.)

SO Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,42e-241	Length:	7310
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	2	Gaps:	0

US-09-930-591-2 (1-686) x AA098221 (1-7310)

```

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB :::
DB 1728 CTGGCGCCCATCAAGGGGTACGCCAGCAAGGGGCTCTTAAAGGTGATATAC 1787
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 1788 AGCTTAATCTGGCGGGGCAAAAACCAAGTGAAGGTATAGTCCAGATTGTGTCACTGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysAlaAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 1848 GCCCAACCTTCTCGGCAACGTGATCAATGGGTGTGTGCTGACTGCTACCAAGGGGCC 1907
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 1908 GGAACGAGGACCAATCGGTACCCCAAGGTCCTGTCTCAGATGTATACCAATGTAGAC 1967
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyValArgSerLeuThrProCysThrCys 100
DB 1968 CAAAGCTTTGTGGCTGGCCCGCTCCGCAAGTACGGCTCATTTGACACCTGCACTTGC 2027
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 2028 GGCTCTCGGACCTTTACTGTGTACAGAGGACGCGCATGTCCCGTGGCGCGCGG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 2088 GGTGATGACAGGGGCGACCTGTCTGCCCGGCCCATTTCTTCTTGAAGGCTCTCTCG 2147
QY 141 GlnGlyProLeuLeuCysProAlaGlnHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 2148 GGGGGTCCGCTGTGTGCCCGGGGCGACCGCGTGGGCATATTAAAGGCGCGGTGCG 2207
QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGlnSerLeuGlnThrThrMet 180
DB 2208 ACCCGTGAAGTGGCTTAAGCGGTGACTTTATCCCTGTGAGAACTTAAGAACACATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 2268 AGGTCCCCCGGTGTTCAAGGATTAACCTCTCCACCAATAGTGGCCGAGAGCTTCAAGTG 2327
QY 201 AlaHisIleuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 2328 GCTCACTCCATGTCTCCACAGGCAAGGCAAAAGCAACAAAGTCCCGGTGCAATAGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 2388 GCTAAGGGCTATTAAGGTGCTAGTACTCAACCCCTGTGTGTGCAACACTGGGCTTTGGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 2448 GCTTACATGTCTCAAGGCTCATGGATGATCTTAATCAAGACCGGGGTGAGAACAAATT 2507

```

```

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 2508 ACCACTGGCAGAGCCCATCAAGTACTCCACTACCGGCAAGTCTTCCGCGCGGGTGC 2567
QY 281 SerGlyGlyAlaLeuThrAspIleIleIleCysAsnGlnCysHisSerThrAspAlaThrSer 300
DB 2568 TCGGGGGGCGCTTATACATTAATTAATTTGAGAGAGTGCACATCCACGGATGCGACATCC 2627
QY 301 IleLeuGlyIleGlyIleThrValLeuAspGlnAlaGlnThrAlaGlnValArgLeuThrVal 320
DB 2628 ATCTTGGGATTCGGCACTGTCTTGAACAGAGAACTCGGGGGCGAGACTGTTGTG 2687
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGlnVal 340
DB 2688 CTCGCCACCGCCACCCCTCGGGCTCCGTCACTGTGCCCATCCCAATCGAGAGGTT 2747
QY 341 AlaLeuSerThrThrGlyGlnIleProPheTyrGlyLysAlaIleProLeuGlnAlaIle 360
DB 2748 GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAAATC 2807
QY 361 LysGlyGlyValArgHisIleuIlePheCysHisSerLysLysLysCysAspGlnLeuAlaAla 380
DB 2808 AAGGGGGGAGACATCTCATCTTCTGTCAATTAAGAAGAGTGCACGAACTCGCGCA 2867
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 2868 AAGCTGGTGCATTTGGGATCAATGCGGTGCTTACCGCGGTCTTGAAGTGTCCGCT 2927
QY 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 2928 ATCCCAACGACCGGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 2988 GGGCACTTGACTCGGTGATGATCTGATCATACGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3047
QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 3048 CTTGACCTTACCTTCAACATTAAGACATCAAGCTCCCGGATGCTGTCTCCGACT 3107
QY 461 GlnArgArgGlyValArgThrGlyValArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 3108 CAAAGTGGGGGACGACTGCGCAAGGGAAGCCAGGATCAAGATTTGTGGACCGGGG 3167
QY 481 GlnArgProSerGlyLysMetPheAspSerSerValLeuCysGlnCysTyrAspAlaGlyCys 500
DB 3168 GAGCGCCCTCGGCATGTTCGACTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3227
QY 501 AlaTrpTyrGlnLeuThrProAlaGlnThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 3228 GCTTGTATGAGCTCAAGCCCGCCGAGACTCAAGTAAAGTAAAGGCTTAAGCAAGCAACC 3287
QY 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheTrpGlnGlyValPheThrGlyLeu 540
DB 3288 CCGGGGCTTCCGTGTGCGAGACCATCTTGAATTTTGGAGGGCGCTTTTAAAGGCGCTC 3347
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrGlyGlnSerGlyGlnAsnLeuProTyr 560
DB 3348 ACTCATTAATATCCCACTTCTATCCCAAGCAAGAGTGGGAGAACTTCTTAC 3407
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB 3408 CTGTAGCGTACCAAGCCACCGTGTGCGCTAAGGCTCAAGCCCTCCCATGTGGGAC 3467
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrIleuHisGlyProThrProLeuLeu 600
DB 3468 CAGATGTGAAGTGTGATTTCCCTCAAGCCCACTCATGTGGCAACACCCGTGCTA 3527
QY 601 TyrArgGlnGlyValAlaGlnAsnGlnValThrLeuThrHisProValThrLysTyrIle 620
DB 3528 TACAGACTGGGCGCTGTTCAGATGAATATCACTTACGACCCAGTCAACCAATATCATC 3587

```

```
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValIleuValGlyGly 640
DB 3588 ATGCATGCGATGTCGGCGGACCTCGAGCTCTCCACGACGACCTGGGCTCTGTTGGCGGC 3647
QY 641 ValIleuAlaAlaIleuAlaAlaIleuValCysLeuSerThrGlyCysValIleuValGlyArg 660
DB 3648 GTCCTGCGCTGCTTGGCGCGGCTATGCTGCAACAGCGCTCGGTCATATGTCGAGG 3707
QY 661 IleValIleuSerGlyLysProAlaIleIleProAspArgGluValIleuValArgGluPhe 680
DB 3708 GTCGCTTGTTCGGGAAAGCCGACATCATCTGACAGGAAAGTCTCTACCGAGAGTTC 3767
QY 681 AspGluMetGluGluCys 686
DB 3768 GATGAGATGGAAGAGTGC 3785

RESULT 7
AAA75296
ID AAA75296 standard; cDNA; 8316 BP.
XX
AC AAA75296;
DT 15-JAN-2001 (first entry)
XX
DE cDNA sequence compiled Hepatitis C virus cDNA clones.
XX
KM Hepatitis C virus; HCV, antisense polynucleotide; polypeptide;
XX viral infectivity; viral replication; ds.
XX
OS Hepatitis C virus.
XX
FH Key 1. 8316
FT CDS /*tag=
FT /note= "partial sequence; no termination codon given"
XX
XX EPI034785-A2.
XX
PD 13-SEP-2000.
XX
PF 16-MAR-1990; 2000EP-00109602.
XX
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341334.
PR 18-MAY-1989; 89US-00355002.
PR 16-MAR-1990; 90EP-00302866.
XX
XX (CHIR ) CHIRON CORP.
XX
PA Houghton M, Choo Q, Kuo G;
PI WPI; 2000-566891/53.
PI P-PSDB; AAB18540.
XX
DR Novel composition comprising a hepatitis C virus antisense polynucleotide
PT which is complementary to or corresponds to a sense strand of the virus
PT genome, and selectively hybridizes to it.
XX
XX Example; Fig 16; 75pp; English.
XX
XX The specification describes a pharmaceutical composition which comprises
XX a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
XX characterized by a positive stranded RNA genome which has 40% homology at
XX the polypeptide level to a HCV polypeptide. The antisense polynucleotide
XX binds to cellular polynucleotides which enhance and/or are required for
XX viral infectivity, replicative ability or chronicity. The antisense
XX polynucleotides may also be designed to bind with high specificity, to be
XX of increased stability, to be stable and to have low toxicity. The
XX composition also comprises an agent which causes viral RNA to be
XX inactive. The composition is used for preventing HCV replication in a
XX system. The present sequence represents a novel HCV cDNA sequence, which
XX is used in the course of the invention
```

```
SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1 66e-241 Length: 8316
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x AAA75296 (1-8316)
QY 1 MetAlaProIleThrAlaTrpAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
DB 2734 CTGGCGCCATGTCACGGCGTACGCGCCAGACAGACAGAGGGCCCTTACGGTGCATATCACC 2793
QY 21 SerLeuThrGlyArgAspLysAlaGlnValGluGlyValGlnIleValSerThrAla 40
DB 2794 AGCCTTAAGTGGCCGGGCAAAAACCAAGTGAAGGGTGAAGTCCAGATTGTCAGACTGCT 2853
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 2854 GCCCAAACTTCTCTGGACAGTCATCATATGGGTGCTGACCTGTACACAGGGGCC 2913
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 2914 GGAACAGAGACCATCGCGTACCAACAGGCTCTGTATCATGATATACCAATGTATGAC 2973
QY 81 GluAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 2974 CAAGACCTGTGTGGCTGGCCGCTCGCAAGTGAAGCTGATTCATTCACCTTCACCTGTC 3033
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 3034 GGCTCTCGACCTTTCCTGTCGTCACAGAGCAGCCGATCATTCCTCCGCGCCGGCGG 3093
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 140
DB 3094 GGTGATAGCAGGGGCGAGCTGTGTGCGCCCGCCCATTTCTTAAAGGCTCTCTCG 3153
QY 141 GlyGlyProLeuLeuLeuProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3154 GGGGCTCGCTGTGTGCCCCGCGGCGAGCGCTGGGCAATATTAAGGCGCGGCTGTGC 3213
QY 161 ThrArgGlyValAlaIleValAlaValAspPheIleProValGlnSerLeuGlnThrMet 180
DB 3214 ACCGTGAGTGGCTAAGGCGGTGACTTATTCCTGTGAGAACCTAGACACATCATG 3273
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3274 AGGTCCCGGTGTTCACGAGTAACTCCTCCACACAGTAGTGCCCGACAGCTTCCAGGTG 3333
QY 201 AlaHisLeuHisAlaProThrArgLysSerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3334 GCTCACTTCATCTCCACAGGACAGGCGGAAAGCAACAGATCCCGGCTGATATGCA 3393
QY 221 AlaGlnGlyTyrLysValIleuValIleuAspProSerValAlaIleThrMetGlyPheGly 240
DB 3394 GCTCAAGGCTATTAAGGTGCTAGTACTCAACCTCTGTCTCTGCAACACCTGGCTTGGT 3453
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 3454 GCTTACATGTCCAAAGGCTCATGGATGATCTTAATCATCGACCGGGGTGAACACAT 3513
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 3514 ACCACTGGAGCCCAACAGTACTCCACTACGCGCAAGTCTTGGCGAGGGGGTGC 3573
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 3574 TCGGGGGGCGCTTATGACATATATTTTGAAGAGGCCACTCCACAGATGCAATCC 3633
QY 301 IleuGluGlyIleGlyThrValIleuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
```

```

Db      3634 ATCTGGGCGATCGGCACCTGCTTGACCAAGCAGAGACTGGGGGGGAGACGTGGTGTG 3693
Qy      321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      3694 CTGGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCACTCAGAGAGATT 3753
Qy      341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      3754 GCTCTGTCCACCAACCGAGAGATCCTTTTACGGCAAGGTATCCCTCCCTGAAAGTATC 3813
Qy      361 LysGlyGlyArgHisLeuIlePheCysHisSerIleLysLysCysAspGluLeuAla 380
Db      3814 AAGGGGGGGAGACATCTCATCTTCTTCATTCAAAGAAAGAGTCGACGAACTGGCCGCA 3873
Qy      381 LysLeuValAlaLeuGlyLysAlaAspAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      3874 AAGCTGGTGGATTTGGGCATCAATGCGGTGGCTTACTACCGGCTTTGACGTGTCCGTC 3993
Qy      401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      3934 ATCCGACCAAGCGCGCATGTGTGTGTGGCAACGATGCCCTCATGACCGGCTATACC 3993
Qy      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      3994 GGCACACTTCGACTCGGTGATGACTGCAATACGTGTGTCAACCCAGACAGTCAATTGACG 4053
Qy      441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      4054 CTTGACCTTACCTTACCATTTGAGACAAATACGGCTCCCCCAAGATGCTGTCTCCGCACT 4113
Qy      461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db      4114 CAACGTGGGGGAGAGACTGGGAGGGGAGGAGCAAGCATTCACGATTTGTGGACCGGGG 4173
Qy      481 GluArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTyrAspAlaGlyCys 500
Db      4174 GAGGCGCCCTCCGCGCATGTTGCACTCGCTCCCTCTGTGAGTGTGACGACGAGCTGT 4233
Qy      501 AlaTyrPyrGluLeuThrProAlaGluThrThrAlaArgLeuAlaTyrMetAspThr 520
Db      4234 GCTTGGTATAGCTCAGCGCCCGGAGACATCACTAGGCTACGAGGGTACATGAACACC 4293
Qy      521 ProGlyLeuProValCysGlnAspHisLeuGlnPheThrGluGlyValPheThrGlyLeu 540
Db      4294 CCGGGGCTTCCCGTGTCCAGAACCATCTTGAAATTTGGAGGGCGCTTTACAGGCGCTC 4353
Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysAsnLeuProTyr 560
Db      4354 ACTCATATAGATGCCCACTTCTTATCCACAGCAAGAGAGTGGGAGAACCTTCTTAC 4413
Qy      561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProPheSerThrAsp 580
Db      4414 CTGGTACCGTACCAAGCACCAGTGTGGCTAGGGCTCAAGGCCCTTCCCACTGTGGAC 4473
Qy      581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      4474 CAGATGTGGAAGTGTGATTCGCTCAAGCCCACTTCATGGGCCAACAACCCCTCTTA 4533
Qy      601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      4534 TACAGACTGGGGCGCTGTTCAAGATGAATACCCCTGACGACCCAGTCAACAATATCATC 4593
Qy      621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db      4594 ATGCAATGACATGTGCGCGGACCTGGAGAGTGTCCGACGACCTGGGTGCTGTGGGGGC 4653
Qy      641 ValLeuAlaIleAlaValAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      4654 GTCTGTGCTGCTTGGCGCGGTATTTGCTGTCAACAGGCTCGGTGTCATATGTGGGAGG 4713
Qy      661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680

```

```

Db      4714 GTCTCTTGTCCGGGAAGCCGCGCAATCATCTGACAGAGGAGAGTCTTACGAGAGTTC 4773
Qy      681 AspGluMetGluGluCys 686
Db      4774 GATGAGATGGAAGATGTC 4791

RESULT 8
AAZ07656
ID  AAZ07656 standard; DNA; 9133 BP.
XX
AC  AAZ07656;
XX
DT  20-MAR-2003 (revised)
DT  08-NOV-1999 (first entry)
XX
DE  Nucleotide sequence of HCV-1 ORF.
XX
KM  Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;
KM  HCV infection; vaccine; ds.
XX
OS  Hepatitis C virus.
XX
FH  Key
FT  CDS
FT  Location/Qualifiers
FT  268..9132
FT  /*tag= b
FT  /transl_except= (pos:1588..1589; aa:Leu)
FT  /note= "this codon has an apparent 1 nucleotide deletion,
FT  which alters the reading frame"
FT  /transl_except= (pos:1647..1650; aa:Pro)
FT  /note= "this codon has an apparent 1 nucleotide
FT  insertion, which alters the reading frame; this insertion
FT  is not indicated in the sequence present in the formal
FT  sequence listing of the specification"
XX
PN  EP939128-A2.
XX
PD  01-SEP-1999.
XX
PF  17-SEP-1990; 99EP-00101746.
XX
PR  15-SEP-1989; 89US-00408045.
PR  21-DEC-1989; 89US-00456142.
PR  17-SEP-1990; 90EP-00310149.
XX
PA  (OYAA/) OYA A.
XX  (CHIR ) CHIRON CORP.
XX
PI  Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;
PI  Cha T, Irvine BD;
XX
DR  WPI; 1999-480843/41.
XX  P-PSDB; AAY14975.
XX
PT  New Hepatitis C Virus isolates, useful for diagnosis of hepatitis
PT  infections and development of vaccines.
XX
PS  Disclosure; Fig 12; 132pp; English.
XX
CC  The invention provides two new isolates of hepatitis C virus (HCV), J1
CC  and J7. These two isolates comprise nucleotide and amino acid sequences
CC  that are distinct from the HCV isolate HCV-1. The nucleotide sequences
CC  may be used to detect non-A, non-B HCV (NANBH) polynucleotides by
CC  hybridisation for diagnosis of NANBH infections. They may also be used to
CC  screen blood donors, donated blood and blood products for this infection.
CC  The isolates may also be used to isolate other naturally occurring
CC  variants of the virus. The polypeptides may be used as a vaccine for
CC  administration to patients to protect against infection with NANBH. The
CC  present sequence represents the nucleotide sequence of HCV-1 ORF.
CC  (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC  correct PR field.)
XX
SQ  Sequence 9133 BP; 1834 A; 2772 C; 2600 G; 1927 T; 0 U; 0 Other;

```


QY 681 AspGluMetGluGluCys 686
 DB 5383 GATGATGGAAGAGTGC 5400

RESULT 9
 AAQ10566
 ID AAQ10566 standard; DNA; 9185 BP.
 XX
 AC AAQ10566;
 XX
 DT 25-MAR-2003 (revised)
 DT 23-APR-1991 (first entry)
 XX
 DE Hepatitis C virus strain 1 DNA.
 XX
 KM Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;
 KM viral infections; ss.
 XX
 OS Hepatitis C virus.
 OS
 PN BP414475-A.
 PN
 PD 27-FEB-1991.
 PD
 XX 21-AUG-1990; 90EP-00309120.
 PF
 XX 25-AUG-1989; 89US-00398667.
 PR
 XX (CHIR) CHIRON CORP.
 PA
 PI Weiner AJ, Steimer KS;
 PI
 DR WPI; 1991-059670/09.
 DR
 XX
 PT Cell lines infected with hepatitis C virus - are used as source of
 PT antigens for detection of HCV antibodies, for vaccines, and for screening
 PT anti-viral agents.
 PS
 XX
 PS Disclosure; Fig 1; 24pp; English.
 XX
 CC This is a hepatitis C virus (HCV) composite cDNA sequence, deduced using
 CC overlapping clones. A compsn. contg. the antigenic protein encoded by
 CC this sequence is useful for detecting anti-HCV anti-bodies (Abs) and for
 CC screening an agent which inhibits HCV replic- action. A cell line infected
 CC with this virus can be used as a source of antigens. The antigen is
 CC useful for preparing vaccines for treating viral infections. See also
 CC AAQ10567. (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 1.88e-241 Length: 9185
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 2 Gaps: 0

US-09-930-591-2 (1-686) x AAQ10566 (1-9185)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysAlaIleThr 20
 DB 3395 CTGGCGCCCATCAGCGGCTAGCGCCACAGCAAGGGGCTCTTCAAGGTGATATACCC 3454

QY 21 SerLeuThrGlyArgAspIysAsnGlnValGluGlyValGlnIleValSerThrAla 40
 DB 3455 AGCTTAAGTGGCCGGGCAAAACCAAGTGGAGGGTCAAGTTCATGTCACACTGCT 3514

QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTerPheValTyrPheGlyAla 60
 DB 3515 GCCCAAACTTCTCCGGCAACCTGCATCATGGGGTGGTGGACTGTCTCAACAGGGGCTC 3574

QY 61 GlyThrArgThrIleAlaSerProIleGlyProValIleGlnMetTyrThrAsnValAsp 80

DB 3575 GGAACAGAGACCATCCGATCACCAAGGGTCCGTGATCCAGATGATACCAATGTAGAC 3634

QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 3635 CAAAGACCTTGTGGGCTGGGCCCTCCGCAAGTACCGCTCATTTGACACCTTCACATTGC 3694

QY 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArg 120
 DB 3695 GGCTCCTGGACCTTTACCTGATGATCAGAGGACGCCGATGTATTCGCCGGCCGGCGG 3754

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlySerSer 140
 DB 3755 GGTGATAGCAGGGGCGACCTGTGTCGCCGCCCATTTCTCACTTGAAAGGCTCCTCG 3814

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 3815 GGGGGTCCGCTGTTGGCCCCCGGGGACAGCCGTGGCATTTTGGGGCCGGCGGTGTC 3874

QY 161 ThrArgGlyValAlaIysAlaValAspPheIleProValGluSerLeuGluThrMet 180
 DB 3875 ACCGTGGAGTGGCTAAGGGGTGAGCTTTATCCTGTGGAGAACCTTAGAGACACATG 3934

QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 3935 AGGTCCCGGTGTTCAAGGATACCTCTCCACAGTAGTGCCCAAGGCTTCAGAGTG 3994

QY 201 AlaHisLeuHisAlaProThrArgIysSerGlyIysSerThrIysValProAlaAlaTyrAla 220
 DB 3995 GCTCACCTCCATGCTCCACAGGACGGGAAAGCAACAGGTCGGGTGCATATGCA 4054

QY 221 AlaGlnGlyTyrIysValLeuValLeuAsnProSerValAlaIleThrMetGlyPheGly 240
 DB 4055 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACATCGGCTTTGGT 4114

QY 241 AlaTyrMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 4115 GCTTACATGCTCAAGGCTCATGGATCATCTCAATCATCAGACCGGGGTAACAACATT 4174

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIysPheLeuAlaAspGlyGlyCys 280
 DB 4175 ACCACGTGGACGCCCATCAGTACCTCACCTACGGGAAGTCTTGCCGACGGCGGGTGC 4234

QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
 DB 4235 TCGGGGGGCGCTTATGATCATTAATTTGTGACGAGTGCACATCCAGATGCCACATCC 4294

QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 4295 ATCTGGGCAATCGGCACTGCTTGAACAAGCAGAGATCGGGGGCGAGACTGTTGTG 4354

QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 4355 CTGGCAACCGCACCCCTCGGGCTCGTACATGTCGCCCATCCAAACATCGGAGAGTT 4414

QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyIysAlaIleProLeuGlnAlaIle 360
 DB 4415 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGGCAAGCTATCCCTCCGAAGTATTC 4474

QY 361 IysGlyIysArgHisLeuIlePheCysHisSerIysAlaIysCysAspGluLeuAlaIle 380
 DB 4475 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAGAGCGACAGCAATCGCGCGCA 4534

QY 381 IysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 4535 AACCTGTGTCATTTGGCATCAATGCGGTGCTTACCGCTTACCGGTCTTACGTCCTGTC 4594

QY 401 IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
 DB 4595 ATCCCAACAGCGGCGATGTTGTCGTGGCAACGATGCCCTCATGACCGGCTATACC 4654

QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrGlnThrValAspPheSer 440


```

Db      4655 GGGGACTTGCATCTCGGATGATGACTGCATACGCTGTCACCCAGACAGTCATTCAGC 4714
Qy      441  LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      4715 CTGACCTTACCTTACCATTTGAGACATACACGCTCCCGAGATGCTGCTCCCGCACT 4774
Qy      461  GlnArgArgGlyArgThrGlyArgGlyValProGlyIleThrArgPheValAlaProGly 480
Db      4775 CAACGTGGGGGCGAGCATGGCGAGGGGAGACCGACATTCACATGATTTGGCGACCGGG 4834
Qy      481  GlnArgProSerGlyMetPheAspSerValLeuGlyGlyCysTyrAspAlaGlyCys 500
Db      4835 GAGGGCCCTCCCGCATGTTGACTGCTCGCTCTGTGATGCTATGACCGCAGGCTGT 4894
Qy      501  AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db      4895 GCTTGATGATGAGCTCAGCGCCCGGACATACATGAGGTACGACGATCATGACAC 4954
Qy      521  ProGlyLeuProValCysGlnAspHisLeuGluThrPheTrpGluGlyValPheThrGlyLeu 540
Db      4955 CCGGGGCTCCCGTGGCGACGACCATCTTCATTTGGAGGGCGCTTTACAGGCTC 5014
Qy      541  ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db      5015 ACTCATATAGATGCGCCACTTTCTATCCACAGCAAGCAGAGTGGGAGAACTTCTTAC 5074
Qy      561  LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db      5075 CTGGTAGGTACCAAGCACCGTGTGGCTCAAGGCTCAAGCCCTCCCGCATCGTGGAC 5134
Qy      581  GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      5135 CAGATGTGAGAGTGTTCATTCGCTCAAGGCCACCTCCCATGGGCCCAACCCCTGCTA 5194
Qy      601  TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      5195 TACAGACTGGGCGCTTCAAGATGAAATCACCTCGAGCGACCCAGTCACCAATATCATC 5254
Qy      621  MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGly 640
Db      5255 ATGACATCATGTCGGCGACCTGAGGTCTGCACAGCACCTGGGTCTGCTGGCGC 5314
Qy      641  ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      5315 GTCCGTGGCTGTTGGCCGCGATTCGCTCAACAGGCTCGTGTATGATGAGTGGCAGG 5374
Qy      661  IleValLeuSerGlyLysProAlaIleIleProAspArgGlyValLeuTyrArgGluPhe 680
Db      5375 GTCCGTCTGTCCGGGAGACCGGCATATCCTGACAGGAAATCTCTTACCGAGAGTTC 5434
Qy      681  AspGluMetGluGluCys 686
Db      5435 GATGAGATGAGAGAGTGC 5452

RESULT 10
ID      AAA75297 standard; cDNA; 9185 BP.
AC      AAA75297;
DT      15-JAN-2001 (first entry)
DE      Sense strand of HCV encoding a polypeptide.
XX      Hepatitis C virus; HCV; antisense polynucleotide; polypeptide;
XX      viral infectivity; viral replication; ds.
OS      Hepatitis C virus.
Key     Location/Qualifiers
FT      CDS          320..9184
FT      /tag=a
FT      /note="partial sequence; no termination codon given"

```

```

XX      XX
PN      EP1034785-A2.
XX      13-SEP-2000.
XX      16-MAR-1990; 2000EP-00109602.
XX      17-MAR-1989; 89US-00325338.
XX      20-APR-1989; 89US-00341334.
XX      18-MAY-1989; 89US-00355002.
XX      16-MAR-1990; 90EP-00302866.
XX      (CHIR ) CHIRON CORP.
XX      Houghton M, Choo Q, Kuo G;
XX      WPI, 2000-566891/53.
XX      P-PSDB; AAB18541.
XX      Novel composition comprising a hepatitis C virus antisense polynucleotide
XX      PT which is complementary to or corresponds to a sense strand of the virus
XX      PT genome, and selectively hybridizes to it.
XX      Example, Fig 17; 75pp; English.
XX      The specification describes a pharmaceutical composition which comprises
XX      CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
XX      CC characterized by a positive stranded RNA genome which has 40% homology at
XX      CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide
XX      CC binds to cellular polynucleotides which enhance and/or are required for
XX      CC viral infectivity, replicative ability or chronicity. The antisense
XX      CC polynucleotides may also be designed to bind with high specificity, to be
XX      CC of increased stability, to be stable and to have low toxicity. The
XX      CC composition also comprises an agent which causes viral RNA to be
XX      CC inactive. The composition is used for preventing HCV replication in a
XX      CC system. The present sequence represents a novel HCV cDNA sequence, which
XX      CC is used in the course of the invention
XX      SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 1.88e-241 Length: 9185
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: Gaps: 0
US-09-930-591-2 (1-686) x AAA75297 (1-9185)

```

```

Qy      1  MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3395 CTGGCGCCCATCATCAGCGGTACGCCACCAAGAGGGGCTCTTAGGGTCAATATCACC 3454
Qy      21  SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db      3455 AGCTTAATGACCGCGGACAAHAAACCAAGTGAAGGTATGATTCAGATTTGTAACGCT 3514
Qy      41  AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db      3515 GCCCAACCTTCTCGGCACATGTCATCATGAGGTGTGCTGATCTGCTACACCGGGGCC 3574
Qy      61  GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrTrpAsnValAsp 80
Db      3575 GGAACGAGGACCATCGGTACCCAGAGGTCTGTCTATCCAGATGTATACCATGTGAC 3634
Qy      81  GlnAspLeuValGlyTrpProAlaProGlnGlyValAlaArgSerLeuThrProCysTrpCys 100
Db      3635 CAAGACTTGTGGCTGGCCGCTCCGCAAGTAGCCGCTCATTTGACACCCCTGCACCTGC 3694
Qy      101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db      3695 GGCTCTCTGAGACTTTACTCTGTGTACAGAGGACCGCCGATGTCAATTCCTGTCGCGCGG 3754

```

```

QY 121 G1YAspG1YArgG1YSerLeuLeuSerProArgProIleSerTYrLeuLYSG1YSerSer 140
DB 3755 GGATGATGACAGGGGACCTGCTGTCCGCCGCCATTCTTCACTTGAAGAGCTCTCG 3814
QY 141 G1YGLYrProLeuLeuCYeProAlaG1YH1SA1AValG1Y1LePheAlaG1A1AValCYe 160
DB 3815 GGGGGTCCGCTGTGTGCCCCCGGGGACGCCCTGGGCATATTATTAAGCGCGCGTGTGC 3874
QY 161 ThrArgG1YValAlaLYSA1AValA1AAspPhe1LeProValG1uSerLeuG1uThrThrMet 180
DB 3875 ACCGTGAGAGTGGCTTAAGCGCGTGAACCTTATTCCTTGGAGAACCTTAAGACACCAATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProG1uSerTYrG1uVal 200
DB 3935 AGGTCCCCGGGTGTTCAACGAGTAACCTCTCCACCAAGAGTGGCCCAAGAGCTTCAAGTGC 3994
QY 201 AlaH1A1eUuH1SA1AProThrG1YSerG1YLYSserThrLYSValProAlaAlaTYrAla 220
DB 3995 GCTCACTCCATGCTCTCCACAGCGCGCAAAAGCCAAAGTCCCGCGCTGCATATGCA 4054
QY 221 AlaG1uG1YTYrLYSValLeuValLeuAsnProSerValAlaAlaThrMetG1YPhG1Y 240
DB 4055 GCTCAGGGCTATTAAGGTAGTACTAACCCTCTGTGTGCAACAACCTGGGCTTGTGT 4114
QY 241 AlaTYrMetSerLYSA1AHSGLY1LeAspProAsn1LeArgThrG1YValArgThr1Le 260
DB 4115 GCTTACATGTCACAGGCTCATGGGATCGATCTTACATCAGACCGGGGGTGAACAAT 4174
QY 261 ThrThrG1YSerPro1LeThrTYrSerThrTYrG1YLYSPhelEuAlaAspG1YLYCYs 280
DB 4175 ACCACTGGCAGCCCATCAGTACTCCACCTAAGGAGTCTTGGCGAGCGGGGTGC 4234
QY 281 SerG1YGLYAlaTYrAsp1Le1Le1CYeAspG1uCYeH1SerThrAspAlaThrSer 300
DB 4235 TCGGGGGGCTTATGACATTAATTTGTGACAGAGTCCACTCCAGAGATGCCACATCC 4294
QY 301 1LeuG1Y1LeG1YThrVal1LeuAspG1uAlaG1uThraG1YAlaArgLYeThrVal 320
DB 4295 ATCTTGGGCATCGGCATGCTCTTGAACCAAGACAGACTCGGGGGGAGACAGTGTGTG 4354
QY 321 LeuAlaThra1aThrProProG1YSerValThrValProH1SProAsn1LeG1uG1uVal 340
DB 4355 CTGGCCACCGCCACCTCCGGGGCTCCGTCACTGTGCCCATCCCAATCCAGAGAGTT 4414
QY 341 AlaLeuSerThrThrG1YGLY1LeProPheTYrG1YLYSAla1LeProLeuG1uAla1Le 360
DB 4415 GCTCTGTCCACCAACGAGAGATCCCTTTTTCAGGCAAGGCTATCCCTCGAAGTATC 4474
QY 361 LYSG1YGLYArgH1SLeu1LePheCYeH1SerLYSA1LYSLYSCYSAAspG1uLeuAla1A 380
DB 4475 AAGGGGGGGAGACATCTCATCTTCTGTCACTTCAAGAGAGTGCAGACATCGCCGCA 4534
QY 381 LYSA1eUVal1A1eUenG1YValAAsnAlaValAlaTYrTYrArgG1YLeuAspValSerVal 400
DB 4535 AAGGTGGTCCATTTGGGCATCAATGCCGTGGCTACTACCGCGGTCTTGAAGTGTCCGT 4594
QY 401 1LeProThrSerG1YAspValValValValAla1aThraAspAlaLeuMetThrG1YPhThr 420
DB 4595 ATCCCGACCAAGCGCGATGTGTGTGTGTGTGCAACCATGCCCTCATGACCGGCTATACC 4654
QY 421 G1YAspPheAspSerVal1LeAspCYeAsnThrCYeVal1ThrG1uThrValAspPheSer 440
DB 4655 GGGCACTTCACTCGGATAGACTGCATACG1GTGTCAACCCAGACAGTGCATTTAGC 4714
QY 441 LeuAspProThrPheThr1LeG1uThr1LeThrLeuProG1uAspAlaVal1SerArgThr 460
DB 4715 CTGACACCTTACCTTCACTTGAAGACATACAGCTTCCCAAGAGATGTGTCTCCGCACT 4774
QY 461 G1uArgArgG1YArgThrG1YArgG1YLYSProG1Y1LeTYrArgPheValAlaProG1Y 480
DB 4775 CAACGTGGGGCAGAGACTGGCAGGGGGAAGCCAGCATCTTAAGATTGTGTGACCGGG 4834

```

```

QY 481 G1uArgProSerG1YMetPheAspSerSerValLeuCYSG1uCYeTYrAspAlaG1YCYs 500
DB 4835 GAGCGCCCTCCCGGCATGTTGCACTGCTGCTGTGTGATGATGACAGAGGCTGT 4894
QY 501 AlaTYrTYrG1uLeuThrProAlaG1uThrThrValArgLeuArgAlaTYrMetAsnThr 520
DB 4895 GCTTGGATGAGTCAACGCCCGCGAGACTACAGTTAGGTACAGACGTACATGAACACC 4954
QY 521 ProG1YLeuProValCYeG1uAspH1SLeuG1uPheThrProG1YVal1PheThrG1YLeu 540
DB 4955 CCGGGGCTTCCCTGTGCGCAGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCTC 5014
QY 541 ThrH1S1LeAspAlaH1SPhelEuSerG1uThrLYSG1uSerG1YGLYAsnLeuProTYr 560
DB 5015 ACTCATATGATGATGCCCATTTTATCCAGACAAGCAGAGTGGGAGAACTTCTCTTAC 5074
QY 561 LeuValAlaTYrTYrG1uThrVal1CYeAlaArgAlaG1uAlaProProProSerTPAsp 580
DB 5075 CTGATGAGCTTACCAACCAACCGGTGCGCTAGGGGCTCAAGCCCTCCCATGTGGGAC 5134
QY 581 G1uMetThrLYSCYSA1eU1LeArgLeuLYSProThrThrLeuH1SG1YProThrProLeuLeu 600
DB 5135 CAGATGTGAAAGTGTGATTGCTGCTCAAGCCCATCGATGGCCACACCCCTGTCTA 5194
QY 601 TYrArgLeuG1YAlaValG1uAsnG1uValThrLeuThrH1SProValThrLYSTYr1Le 620
DB 5195 TACAGACTGGGCGCTGTGAGATGAATACCTCGACCGCACCCAGTCAACCAATATCATC 5254
QY 621 MetThrCYeMetSerAlaAspLeuG1uValValThrSerThrTYrValLeuValG1YArg 640
DB 5255 ATGACATGATGATGTGGCGCAGCTGAGAGTGTGTCAACAGACCTCGGGTGTCTGTGGGGC 5314
QY 641 ValLeuAla1aLeuAla1aTYrCYeLeuSerThrG1YCYeValVal1LeValG1YArg 660
DB 5315 GTCTGCTGCTTGTGGCGGATGTCTGTCAACAGCGCTGCTGATAGTGGGAGG 5374
QY 661 1LeValLeuSerG1YLYSProAla1Le1LeProAspArgG1uValLeuTYrArgG1uPhe 680
DB 5375 GTGCTGTGTGCGGAAAGCGGCATCATACCTGACAGAGGAATCCTTACGAGAGTTCC 5434
QY 681 AspG1uMetG1uG1uCYs 686
DB 5435 GATGAGATGGAAGAGTGC 5452

RESULT 11
AAT12710
ID AAT12710 standard; cDNA, 9401 BP.
XX
AC AAT12710;
XX
DT 25-MAR-2003 (revised)
DT 15-MAY-1996 (first entry)
XX
DE Hepatitis C virus polypeptide.
XX
KW Non-A non-B hepatitis virus; NANBH; HCV; antigen; detection; diagnosis;
antibodies; ds.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..9378
FT /+tag= a
XX
XX EB693687-A1.
XX
XX 24-JAN-1996.
XX
XX 03-APR-1991; 95EP-00114016.
XX
XX 04-APR-1990; 90US-00504352.
XX
PA (CHIR ) CHIRON CORP.

```

XX Houghton M, Choo Q, Kuo G;
XX WPI, 1996-117956/13.
DR P-PSDB; AAR90931.
XX
PT Combinations of synthetic Hepatitis C Virus antigens - provide more
PT effective diagnosis of Non-A, Non-B Hepatitis.
XX
PS Disclosure: Fig 1(A-Y); 53pp; English.
XX
CC The combination comprises an HCV antigen from the C domain (pref. C22 -
CC AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -
CC AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5
CC (AAR90934) domain. The antigens may in the form of a fusion protein, a
CC simple physical mixture, or the individual antigens commonly bound to a
CC solid matrix. They are pref. prepd. by recombinant DNA techniques
CC (primers are given in AAT12711-T12716), but can be synthesised or
CC isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003
CC to correct PF field.)
XX

SO Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,93e-241	Length:	9401
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
		Gaps:	0

US-09-930-591-2 (1-686) x AAT12710 (1-9401)

```
Qy 1 MetLapProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGCGCCATCAGCGGCTAGCCGACAGCAAGGGCCCTCTAGAGGTGATATACCC 3476
Qy 21 SerLeuThrGlyArgAspGlyAsnGlnValGlnGlyGlnValGlnIleValSerThrAla 40
Db 3477 AGCTTAATGCGCGGAGCAAAACCAAGTGAAGGTGAGTCCAGATTGTCTCAACGTCT 3536
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysThrThrValTyrHisGlyAla 60
Db 3537 GCCCAAACTTCCCTGGCAACCTGCATCATGCGGTGCTGAGTCTCTACACAGGGGCC 3596
Qy 61 GlyThrArgThrIleAlaSerProGlySerGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAACGAGACCATCGCTCACCAAGGCTCTGTCTCATCCAGATGATACCAATGTAGAC 3656
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTGTGGCTGGCCGCTCCGCAAGTGAAGCGCTCATTGACACCTGCACATTGC 3716
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3717 GGCTCCCTCGGACCTTACTGTGTCAAGGAGGACCGCATGTCATTCCGTCGCGCGCG 3776
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlyGlySerSer 140
Db 3777 GGTATAGCAGGGGAGGCTCTGTCTGCGCCGCGCATTTCTTCTTGAAGAGGCTCTCG 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGATCCGCTGTGTGCGCCGCGGGGACCGCGTGGGCATATTAAAGGCGCGGTGTC 3896
Qy 161 ThrArgGlyValAlaIleValAlaValAspPheIleProValGlnSerLeuGlnThrThMet 180
Db 3897 ACCGTGAGTGGCTTAAGCGGTGACTTATCTCTGTGGAGAACTTAAGAACACCATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGGTCCCGGTGTTCAAGGATTAATCTCTCTCCACCAAGTATGTCGCCAGAGCTTCAAGTG 4016
```

```
Qy 201 AlaHisLeuHisAlaProThrArgIleSerGlyIleSerThrIleValProAlaAlaTyrAla 220
Db 4017 GCTCACCTCATCTCTCCACAGGACGCGCAAAAGCAACCAAGGTCCCGGTGATATGCA 4076
Qy 221 AlaGlnGlyTyrTyrValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4077 GCTCAGGGCTATAGGTGCTAGTACACCCCTCTGTGTCTGCAACACGCGGCTTGTGT 4136
Qy 241 AlaTyrMetSerIleHisAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4137 GCTTACATGTCACAGGCTCATGGGATCGATTCCTAATCATGACGCGGGGTGAACAATT 4196
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIlePheLeuAlaAspGlyGlyCys 280
Db 4197 ACCACTGGCAGCCCATCAAGTCTCCACCTTACGGAAGTCTTCCGCGCGGGGTGC 4256
Qy 281 SerGlyGlyValArgThrIleIleIleCysAspGlnCysHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGGGCGCTTATAGCATTAATTAATTGACGAGTGCACATCCAGATGCCACATCC 4316
Qy 301 IleLeuGlyIleGlyThrValIleAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCAATCGGCACTGCTCTTGAACAGACAGACATGCGGGGAGACTGTTGTG 4376
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGlnVal 340
Db 4377 CTGGCCACCCGCAACCCCTCCGGGCTCCGTCACATGTGCCCCATCCAAATCGAGAGGT 4436
Qy 341 AlaLeuSerThrThrArgIleGlnIleProPheTyrGlyValAlaIleProLeuGlnAlaIle 360
Db 4437 GCTCTGTCACACCGAGAGATCCCTTTTACGGAGGCTATCCCTCGAAGTATATC 4496
Qy 361 IysGlyGlyValArgHisLeuIlePheCysHisSerIleIleCysAspGlnLeuAlaIle 380
Db 4497 AAGGGGGGAGACATCTCATCTTCTGTGATTAAGAAAGAGTGCAGCAACTCGCGCA 4556
Qy 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4557 AAGCTGTCCCATTTGGGCAATCAATGCGGTGGCTTACTACCGCGTCTTGACGTGTCTGC 4616
Qy 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4617 ATCCGACACGCGCGCATGTTGTGCTGTGCAACGATGCCCTTCATGACCGGCTATACC 4676
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4677 GGGCACTTGACCTCGGATGATGACCTGCATATACGTGTGTCACCGACAGTCAATTACG 4736
Qy 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4737 CTTGACCTTACCTTACCATTTGAGACAAATACGCTCCCGAGATGCTGTCTCCGCACT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyIleProGlyIleTyrArgPheValAlaProGly 480
Db 4797 CAAGTGGGGGAGAGCTGCAAGGAGGAGGAGCAGGATCTACAGATTTGTGGACCGGG 4856
Qy 481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyGlyCysTyrAspAlaGlyCys 500
Db 4857 GAGGCGCCCTCCGAGCATGTTGCACTCGTCCGTCTGTGATGATGACGAGGCTGT 4916
Qy 501 AlaTyrTyrGlnLeuThrProAlaGlnThrThrValArgLeuArgAlaTyrMetLeuThr 520
Db 4917 GCTTGATATAGCTTACAGCGCGCGAGACTACAGTTAGGTTAGAGGTTACATGAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheThrGlnGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTTCCGTGTGCGCAGAGACATCTTGAAATTTTGGAGGGCGCTTTTACAGGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyGlnLeuLeuProTyr 560
Db 5037 ACTCATATATAGTCCGCACTTCTATCCACAGAAAGAGTGGGAGAACTTCTCCATAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
```

```

Db      5097 CTGTAAGCCGTAACAGCAGCAGCTGTGCTAGAGGCTTCAAGCCCTCCCATCGTGGAG 5156
Qy      581  GIMeETThpYsCysLeuIleAArgLeuLysProThrluHsIsgLYProThrProLeuLeu 600
Db      5157 CAGATGTGGAAGGTTTGAATTCGCTTCAAGCCACCTTCAATGGGCCAACACCCCTGCTA 5216
Qy      601  TTAArgLeuGlyAlaValGlnAsnGluValThrLeuThrlsPProValThrLYsTYrIle 620
Db      5217 TACAGACTGGGGCGGTGTTCAAGATGAATCATCCCTGACGCCACCAAGTCAACAAATATCATC 5276
Qy      621  MetThrCYMeSerAlaAspLeuGluValAlaThrSerThrTYrValLeuValGlyGly 640
Db      5277 ATACATGATGATGTCGGCGGACCTGGAGAGTGTCTACGAGCACTGGGGTGTCTGTGGCGGC 5336
Qy      641  ValLeuAlaAlaLeuAlaAlaTYrCYLeuSerThrGlyCYsValAlaIleValGlyArg 660
Db      5337 GTCTGTGCTCTTGTGGCGGGTATGCTGTCAACAGGCTGCGGTGATAGTGGGAGG 5396
Qy      661  IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTYrArgGluPhe 680
Db      5397 GTGCTCTTGTCCGGGAAGCCGGCATCATCTGACAGGAAGTCTTACCGAGAGTTC 5456
Qy      681  AspGluMeGluGluCYs 686
Db      5457 GATGATGATGAGAGTGC 5474

RESULT 12
AAT99981
ID      AAT99981 standard; DNA; 9401 BP.
AC      AAT99981;
XX
XX      25-MAR-2003 (revised)
DT      16-MAR-1998 (first entry)
XX
XX      HCV polyprotein coding sequence.
DE
XX      PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KM      C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW      NS4; de.
XX
XX      Hepatitis C virus.
OS
XX
XX      Key      Location/Qualifiers
FT      CDS      342..9377
FT      FT      /*tag= a
XX
XX      US5683864-A.
PN
XX      04-NOV-1997.
PD
XX      07-JUL-1992; 92US-00910760.
PE
XX
XX      18-NOV-1987; 87US-00122714.
PR      30-DEC-1987; 87US-00139886.
PR      26-FEB-1988; 88US-00161072.
PR      06-MAY-1988; 88US-00191263.
PR      26-OCT-1988; 88US-00263584.
PR      14-NOV-1988; 88US-00271450.
PR      17-MAR-1989; 89US-00325338.
PR      20-APR-1989; 89US-00341334.
PR      21-APR-1989; 89US-00353896.
PR      18-MAY-1989; 89US-00355002.
PR      04-APR-1990; 90US-00504552.
XX
XX      (CHIR ) CHIRON CORP.
PA
XX      Kuo G, Houghton M, Choo Q;
PI
XX      WPI, 1997-548976/50.
DR      P-PSDB; AAM34480.
XX

```

```

PT      Combination of three hepatitis C virus antigens - used for detection of
XX      specific antibodies to diagnose infection.
XX
XX      Disclosure; Col 25-46; 57pp; English.
CC      This sequence represents the Hepatitis C virus polyprotein coding
CC      sequence. Fragments of this sequence can be amplified and used in the
CC      combination of HCV antigens of the invention. The HCV antigen combination
CC      comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
CC      1-120 of the HCV polyprotein), or its immunologically reactive fragment
CC      containing at least 8 aa. It also comprises two additional antigens from
CC      two different polyprotein domains, including at least 8 aa from the NS3,
CC      NS4, S or NS5 domains of the polyprotein, corresponding, respectively, to
CC      aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein.
CC      Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
CC      regions of the HCV polyprotein. These antigen combinations are used
CC      diagnostically to detect anti-HCV antibodies, using any standard
CC      immunosassay format. These antigen combinations have a broader range of
CC      reactivity with antibodies than any antigen individually. (Updated on 25-
CC      MAR-2003 to correct PR field.)
XX
SQ      Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1,93e-241      Length:      9401
Score:      3574.00      Matches:      672
Percent Similarity:      99.42%      Conservative:      10
Best Local Similarity:      97.96%      Mismatches:      4
Query Match:      98.78%      Indels:      0
DB:      2      Gaps:      0

US-09-930-591-2 (1-686) x AAT99981 (1-9401)
Qy      1      MetAlaProIleThrlaTYrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3417 CTGGCCCATCATCCGCGGTATGCGCCGACGACCAAGGGGCTCTTCAAGTGAATCACC 3476
Qy      21      SerLeuThrGlyARgsApLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      3477 AGCTTAACGTGCGCGGACAAACAAACAAAGTGAAGGTGAGGTCCAGATTGTGTAACGTCT 3536
Qy      41      AlaGlnThrPheLeuAlaThrCYsIleAsnGlyValCYsTrpThrValTYrHsIsgLYAla 60
Db      3537 GCCCAAACTTCTCTGCAACAGTGCATCAATGGGGTGTCTGACGTCTTACCAACCGGGCC 3596
Qy      61      GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTYrThrAsnValaAsp 80
Db      3597 GGAACGAGGACATCGCGTCAACCAAGGATCTGTATCCAGATGATACCAATGATAGAC 3656
Qy      81      GlnAspLeuValGlyTYrTrpProAlaProGlnGlyAlaArgSerLeuThrProCYsThrCYs 100
Db      3657 CAGACCTTGTGGGCTGGCCCGCTCCGCAAGTAGCCGCTCATTTGACACCTGCACTTGC 3716
Qy      101      GlySerSerAspLeuTYrLeuValThrArgHsIAspValIleProValaArgArg 120
Db      3717 GGCTCTCGGACCTTACTTACCTGATGACAGAGCAGCCGATGTATTCCTCGGCGCGGGCGG 3776
Qy      121      GlyAspGlyARgLYserLeuLeuSerProArgProIleSerTYrLeuLysGlySerSer 140
Db      3777 GGTGATAGCAGGGGCGCTGTGCTGCCCGGCCCATTTCTCACTGAAAGGCTCTCTCG 3836
Qy      141      GlyLYProLeuLeuCYsProAlaGlyHsIAspValGlyIlePheArgAlaAlaValCYs 160
Db      3837 GGGGGTCCGCTGTGTGGCCCGCGGGGACGCGGTGGCATATTATTTAGGCGCGGGTGTGC 3896
Qy      161      ThrArgGlyValAlaValaValaAspPheIleProValGlySerLeuGlnThrMet 180
Db      3897 ACCGTGAGTGGCTAAGCGGGGACTTATCTCTGTGAGAAACCTTAAAGAACCAACATG 3956
Qy      181      ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTYrGlnVal 200
Db      3957 AGGTCCCGGATTTCAAGATTAATCTCTCTCAACGATAGTGCACGAGCTTCAGGTTCCAGGTT 4016

```

```

OY 201 AlAHISleuHIsalApProthrglySerGlylySerSerThryValProHlaAlaTyraLa 220
DB 4017 GCTCACCCTTCATGCTCCACAGAGCGAGCAAAAGCAACCAAGGTCCCGGCTGATAGCA 4076
OY 221 AlAGInglYTyrylsValleuValleuAsnProSerValAlaAlaThrmecglyPhegly 240
DB 4077 GCTCAGGGGCTTAAGGTGCTAGTACTCAACCCCTCTGCTGCTGCAACACCGGGCTTGGT 4136
OY 241 AlaTyMeSerlySalahIsGlyleAspProAsnIleArgThrglyValaGlyThrIle 260
DB 4137 GCTTACATGTCACAGGCTCATGGAGATCGATCTTAACATCAGAGCCGGGGGTGAACAACATT 4196
OY 261 ThrThrglySerProIleThrySerThryGlylyPheleuAlaAspGlyGlyCys 280
DB 4197 ACCACTGACAGCCCATCAGACTCCACCTACGCGCAAGTCTTGGCCGACGGGGGTGC 4256
OY 281 SerGlyGlyAlaTyraPheleleleleCysAspGlyCysHisSerThraPalaThrs 300
DB 4257 TCGGGGGGGCGCTTATGACATATATTTGTGACGAGTGCACCTCCAGATGCCACATCC 4316
OY 301 IleleuglylleGlyThryValleuAspGlnAlaGluThraAlaGlyValaArgleuThryVal 320
DB 4317 ATCTTGGGCGATCGGACATGCTCTTGACCAAGACAGACTGGCGGGGCGAGACTGGTTGG 4376
OY 321 LeuAlaThraAlaThraProProGlySerValThryValProHisProAsnIleGluGluVal 340
DB 4377 CTGCGCACCGGCAACCCCTCCGGGCTCCGTCACCTGTGCCCATCTCCCAACATCGAGAGATT 4436
OY 341 AlaIeuserThrrhrglyGluIleProPheThryGlylySalalleProleuGluAlaIle 360
DB 4437 GCTCTGTCCACCAACGAGAGATCCCTTTTACGCGCAAGGCTATCCCTCCGAAAGTATC 4496
OY 361 LysGlyGlyAlaArgHisleuIlePheCysHisSerlyslsCysAspGlnLeuAlaAla 380
DB 4497 AAGGGGGGAGACATCTATCTCTGTCATTAAGAAAGAGTGCGACGAACTGCCGCA 4556
OY 381 LysLeuValAlaLeuGlyValaAsnAlaValAlaTyryraGlyLeuAspValSerVal 400
DB 4557 AAGCTGTCGCGATTTGGGATCATAGCGGTGCTACTACCCGCGCTTGAAGTCTCGCTC 4616
OY 401 IleProthrseryAspValaValaValaValaAlaThraPalaIleuMetThrglyPheThr 420
DB 4617 ATCCCGACGAGCGGAGTGTGTCGTGCGCAACCGATGCGCTCATGACCGGCTATACC 4676
OY 421 GlyAspPheAspSerValIleAspCysAsnThryCysValIThrgInThryValaAspPheSer 440
DB 4677 GCGGACTTCGACTCGGTATGACTGCAATAGTGTGTGACCCGACGAGCTGATTTGAC 4736
OY 441 LeuAspProthrrPheThrrIleGluThrrIleThryleuProGlnAspAlaValaSerArgThr 460
DB 4737 CTTGACCCCTACCTTACCATTTAGACAATACGCTCCCCAGGATGCTGTCTCCGCACT 4796
OY 461 GlnaGlyGlyAlaArgThrglyAlaArgGlylyPheProGlylyleTyraArgPheValAlaProGly 480
DB 4797 CAACGTCGGGAGAGACTGGAGGGGAGAGCCAGGACATTCAGATTTGTGACCCGAGG 4856
OY 481 GluArgProSerGlyMeCPheAspSerSerValleuCysGlyCysTyraAspAlaGlyCys 500
DB 4857 GAGCGCCCTCCGCGATGTGACCTGTCGCTCTGTGATGCTTATGACGACGAGCTGT 4916
OY 501 AlaTyryGluIleuThrrProAlaGluThrrThryValaGleuAlaGlyTyMeLsnThr 520
DB 4917 GCTTGGTATGAGCTACGCGCCCGAGACTACAGTTAGGCTTACGAGCTTACATGAACCC 4976
OY 521 ProGlyleuProValCysGlnAspHisleuGluPheThrglyGlyValaPheThrglyleu 540
DB 4977 CCGGGGGCTTCCGCTGGCGAGACATCTTGAATTTGGGAGGGGCTCTTTACAGGCTTC 5036
OY 541 ThrHisIleAspAlaHisPheleuSerGlnThryGlnSerGlylyleuAsnleuProTy 560
DB 5037 ACTCATATAGATGCCCACTTTCTATCCAGACAAAGACAGATGCGGAGAACTTCTTAC 5096
OY 561 LeuValAlaTyryGlnAlaThryValCysAlaArgAlaGlnAlaProProProSerTrpasp 580

```

```

DB 5097 CTGTAGAGGTATCAAGCACCGCTGCGCTAGGGCTCAAGCCCTCCCATCGTGGAGC 5156
OY 581 GlnMetTrpIysCysLeuIleArgleuIysProThrleuHIsGlyProThrrProleu 600
DB 5157 CAGATGTGAAGTGTGATTTGCTTCCCTCAGAGCCACCTTCATGGGCAACACCTGTCTA 5216
OY 601 TyraGlyGlyAlaValaGlnAsnGluValThryleuThrrHisProValThryIle 620
DB 5217 TACAGACTGGGCGCTTTCAGATGAATACCTTCAGACCGACCCAGTCAACAAATACATC 5276
OY 621 MetThryCysMetSerAlaAspLeuGluValaIleThrsThryValleuValGly 640
DB 5277 ATGACATGATGTGCGCGACCTGAGGTCGTACAGAGACCTGGAGTCTGCTGTGGGCG 5336
OY 641 ValleuAlaAlaLeuAlaAlaTyryCysleuserThrglyCysValaIleValaGlyArg 660
DB 5337 GTCTTGCTGCTTTGGCCCGCGATTCCTGTCAACAGGCTGTGGTATTAAGTGGGACAG 5396
OY 661 IleValleuSerGlylyAspProAlaIleIleProAspArgGluValleuTyraGlyPhe 680
DB 5397 GTGCTTGTTCGGGAAAGCGGCAATCATACCTGACAGGAAATCCTTACCGAGATTCC 5456
OY 681 AspGluMetGluGluCys 686
DB 5457 GATGAGATGGAAGATGTC 5474

RESULT 13
AAV09989 standard; DNA; 9401 BP.
ID AAV09989;
AC AAV09989;
XX 26-MAY-1998 (first entry)
DT HCV polyprotein coding sequence DNA.
DE HCV polyprotein coding sequence DNA.
XX Hepatitis C virus C domain; HCV, C antigen; immunological activity;
KW NS3 domain; NS4 domain; S domain; NS5 domain; de.
XX Hepatitis C virus.
OS
FH Key Location/Qualifiers
FT CDS 342..9377
FT FT /*tag= a
FT FT /product= "HCV polyprotein"
PN US5712087-A.
XX 27-JAN-1998.
PD 12-MAY-1995; 95US-00440519.
XX 04-APR-1990; 90US-00504352.
PR 07-JUL-1992; 92US-00910760.
XX (CHIR ) CHIRON CORP.
PA Kuo G, Houghton M, Choo Q;
XX WPI: 1998-119973/11.
DR P-PSDB: MAM40038.
XX Immunassays for hepatitis C virus antibodies - using combinations of
PT antigenic fragments of HCV polyprotein.
PS Disclosure; Fig 1; 59pp; English.
XX This sequence encodes the hepatitis C virus (HCV) polyprotein which is
CC used in the construction of novel combinations of HCV antigens that have
CC a broader range of immunological activity than any single HCV antigen. An
CC example of such an antigen given in this specification comprises a first
CC antigen containing at least 8 amino acids of the C domain of the HCV

```

CC polypeptide and a second antigen comprising at least 8 amino acids of the
CC NS3 domain, the NS4 domain, the S domain or the NS5 domain of the HCV
CC polypeptide in the form of a fusion protein, a physical mixture or bound
XX to a solid matrix

SO Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,936-241	Length:	9401
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	2	Gaps:	0

US-09-930-591-2 (1-686) x AA09989 (1-9401)

```
QY 1 MetAlaProIleThraIaTYRAlaGlnGlnThraArgGlyLeuLeuGlyCysIleIleThr 20
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3417 CTGGCGCCCATCAAGCGGTACGCCACAGACAAAGGGCCCTCTAGGGGTGATATCAC 3476

QY 21 SerLeuThrGlyArgAspIysAsnGlnValGlnGlyGluValGlnIleValSerThrAla 40
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3477 AGCTTAATGCGCCGGGACAAACCAAGTGAAGGTGAAGTCCAGATTGTCTCAACTGCT 3536

QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTYRHisGlyAla 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3537 GCCCAAACTTCTCCGGACAGCTGATCAATGGGTGTGCTGATGCTGTACACAGGGGCC 3596

QY 61 GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTYRThrAsnValAsp 80
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3597 GGAACGAGAGACATCGCTGACCCCAAGGGTCTGTCTCATCCAGATGTATACCAATGTAGAC 3656

QY 81 GluAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3657 CAACACCTTGAGGCTGCGCCGCTCCGCAAGTAGCGCTCATTTAGACACCTTGACATTGC 3716

QY 101 GlySerSerAspLeuTYRLeuValThraGHisAlaAspValIleProValArgArgArg 120
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3717 GGCCTCTCGACCTTTACTGTGTACAGAGGACGCCCATGTCTATCCCGTGGCCGGCGG 3776

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTYRLeuLysGlySerSer 140
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3777 GGTGATAGACAGGGGACAGCTGTCTGCGCCCGCCATTCTTCTTAAGAAAGGCTCTCG 3836

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3837 GGGGGTCCGCTGTGTGCGCCGCGGGACCGCTGGGCAATATTAGGGCGCGCTGTGC 3896

QY 161 ThrArgGlyValAlaIysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3897 ACCCGTGAAGTGGTTAAGCGCGTGAATTTATCCCTGTGAGAACTTAAGACACCATG 3956

QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTYRLeuVal 200
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3957 AGGCCCCCGGTTCACGAGTAATCTCTTCCACAGAGTGGCCCGAGCTTCAAGTGTG 4016

QY 201 AlaHisIleuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTYRAla 220
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4017 GCTCACTTCATGCTCCACAGGACGGGCAAAAGACCAAGGTCGCCGTGCTCATAGCA 4076

QY 221 AlaGlnGlyTYRLeuValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4077 GCTCAGGGCTTAAGAGTGTAGTACTCAACCCCTCTGTGTGTCACACAGGGCTTTGT 4136

QY 241 AlaTYRMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4137 GCTTACATGTCCAAAGGCTCATGGGATCGATCTTAACATCAGAGCCGGGGTGAAGACATT 4196

QY 261 ThrThrGlySerProIleThrTYRSetThrTYRGLYSPheLeuAlaAspGlyValCys 280
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4197 ACCACTGGACGCCCATCAAGTACTCCACCTTAAGCAAGTTCTTGGCGAGGGGTGTC 4256
```

```
QY 281 SerGlyValAlaTYRAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4257 TCGGGGGGCGCTTATGACATTAATTTGAGAGAGGCCACTCCACGATGCCACATCC 4316

QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4317 ATCTTGGGCAATCGCACTGTCTTGAACAGAGAGACTGCGGGGCGAATCGTTGTG 4376

QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGluVal 340
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4377 CTGCGCACCGCACCCCTCGGGCTCGGTACTGTGCCATCCCAATCGAGAGGATT 4436

QY 341 AlaLeuSerThrThrGlyGluIleProPheTYRGLYSPheValIleProLeuGluAlaIle 360
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4437 GCTGTGTCACCAACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTATC 4496

QY 361 LysGlyGlyArgHisIleuIlePheCysHisSerLysLysLysCysAspGlyLeuAlaAla 380
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4497 AAGGGGGGAGACATCTCATCTTCTGTCAATAAGAGAGCGACAGAACTCGCGCA 4556

QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTYRTrpArgGlyLeuAspValSerVal 400
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4557 AAGCTGTGCAATTGGGCATCAATGCCGTGCTTACTACCGCGTCTTGAAGTCTGCTC 4616

QY 401 IleProThrSerGlyAspValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4617 ATCCGACACAGGGCGCATGTGTGTGTGGCAACGATCCCTCATGACGGCTATAC 4676

QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIleGlnThrValAspPheSer 440
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4677 GGGACCTTCACCTCGGTGATAGCTCAATACCTGTGTCAACCCAGACAGTCAATTCAGC 4736

QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4737 CTTGACCTTACCTTCAACCATGAGACATACCTCCCGAGATGTGTCTCCGACT 4796

QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTYRArgPheValAlaProGly 480
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4797 CAACGTGCGGGGACAGCATGGGCAAGGGGAAAGCCAGCATCATGATTTGTGGACCGGG 4856

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTYRAspAlaGlyCys 500
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4857 GAGCGCCCTCCCGGCAATGTTCGACTGTCCGTCTGTGAGAGTCATGACGAGGCTGT 4916

QY 501 AlaTrpTYRgluLeuThrProAlaGluThrThrValArgLeuArgAlaTYRMetAsnThr 520
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4917 GCTTGTATGAGCTCACGCCCGCGAGACTACGTTAGGCTACAGAGTACATGAACACC 4976

QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4977 CCGGGGCTTCCCTGTGCGCAGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCTC 5036

QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysLeuProTYR 560
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5037 ACTCATATGATATCCCACTTTCATCCACAGCAAAACAGAGTGGGAGAACTTCTTCAC 5096

QY 561 LeuValAlaTYRglnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5097 CTGTGTAGCTTACCAACCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCTCCATGTGGAG 5156

QY 581 GluMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5157 CAGATGTGAAGTGTGATTTGCTCTCAAGCCCACTTCATGGGCAACACCCCTGCTGA 5216

QY 601 TYRArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrLysTYRLe 620
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5217 TACAGACTGGGCGCTGTCAAGATGAATCAACCTGACCAACCACTCAACCAATATCATC 5276

QY 621 MetThrCysMetSerLysLeuGluValValThrSerThrTYRValLeuValGlyGly 640
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5277 ATGACATGATGTGTGGCGCACCTGAGAGTGTACAGACACCTGGGTGCTGTGGCGGG 5336

QY 641 ValLeuAlaAlaLeuAlaAlaTYRcysLeuSerThrGlyCysValValIleValGlyArg 660
```

```

Db      5337  GTCTGGCTGCTTGGCGCGCTATTGCTGCACAGCGCTGCGTGCATATAGTGGCAGG 5396
Qy      661  ILleValIeuSerGlySerProAlaIleIleProAspArgGlyValIleuTyrArgGluPhe 680
Db      5397  GTCTGCTTGTCCGGGAAAGCCGGCAATCACTTAACAGGAAGTCTTACCGGAGATTG 5456
Qy      681  AspGluMetGluGluCys 686
Db      5457  GATGAGATGGAAGAGTGC 5474

RESULT 14
AAD35043
ID      AAD35043 standard; cDNA; 9401 BP.
XX
XX      AAD35043;
XX
XX      16-JUL-2002 (first entry)
XX
XX      Hepatitis C virus (HCV) polypeptide cDNA.
XX
XX      Hepatitis C virus; HCV; antigen; C domain; polypeptide; NS3 domain;
XX      NS4 domain; S domain; NS5 domain; gene; ss.
XX
XX      Hepatitis C virus.
XX
XX      Key
XX      CDS      Location/Qualifiers
XX              342..9377
XX              /*tag= a
XX              /product= "HCV polypeptide"
XX              /transl_except= (pos:366..368, aa:Xaa)
XX              /note= "Xaa equals Lys or Arg"
XX              /transl_except= (pos:372..374, aa:Xaa)
XX              /note= "Xaa equals Thr or Asn"
XX              /transl_except= (pos:867..869, aa:Xaa)
XX              /note= "Xaa equals Thr or Ile"
XX              /transl_except= (pos:1341..1343, aa:Xaa)
XX              /note= "Xaa equals Val or Met"
XX              /transl_except= (pos:2148..2150, aa:Xaa)
XX              /note= "Xaa equals Ile or Leu"
XX              /transl_except= (pos:2883..2885, aa:Xaa)
XX              /note= "Xaa equals Asn or Tyr"
XX              /transl_except= (pos:3681..3683, aa:Xaa)
XX              /note= "Xaa equals Ser or Pro"
XX              /transl_except= (pos:3690..3692, aa:Xaa)
XX              /note= "Xaa equals Thr or Ser"
XX              /transl_except= (pos:4167..4169, aa:Xaa)
XX              /note= "Xaa equals Leu or Pro"
XX              /transl_except= (pos:4323..4325, aa:Xaa)
XX              /note= "Xaa equals Val or Gly"
XX              /transl_except= (pos:4701..4703, aa:Xaa)
XX              /note= "Xaa equals Tyr or Cys"
XX              /transl_except= (pos:4752..4754, aa:Xaa)
XX              /note= "Xaa equals Ser or Thr"
XX              /transl_except= (pos:5970..5972, aa:Xaa)
XX              /note= "Xaa equals Gly or Glu"
XX              /transl_except= (pos:6183..6185, aa:Xaa)
XX              /note= "Xaa equals His or Leu"
XX              /transl_except= (pos:6186..6188, aa:Xaa)
XX              /note= "Xaa equals Cys or Ser"
XX              /transl_except= (pos:6402..6404, aa:Xaa)
XX              /note= "Xaa equals Gly or Val"
XX              /transl_except= (pos:7386..7388, aa:Xaa)
XX              /note= "Xaa equals Ser or Thr"
XX              /transl_except= (pos:7494..7496, aa:Xaa)
XX              /note= "Xaa equals Phe or Tyr"
XX              /transl_except= (pos:7497..7499, aa:Xaa)
XX              /note= "Xaa equals Ala or Ser"
XX              /transl_except= (pos:7845..7847, aa:Xaa)
XX              /note= "Xaa equals Phe or Leu"
XX              /transl_except= (pos:8409..8411, aa:Xaa)
XX              /note= "Xaa equals Gly or Arg"
XX              /transl_except= (pos:9102..9104, aa:Xaa)

```

```

FT      FT      /note= "Xaa equals Gly or Arg"
FT      FT      /transl_except= (pos:9327..9329, aa:Xaa)
FT      FT      /note= "Xaa equals pro or leu; these translational
FT      FT      exceptions occur while decoding for the alternative
FT      FT      version of HCV polypeptide (AAE22052)"
XX      XX      US6312889-B1.
XX      XX      06-NOV-2001.
XX      XX      12-MAY-1995; 95US-00440549.
XX      XX      04-APR-1990; 90US-00504352.
XX      XX      07-JUL-1992; 92US-00910760.
XX      XX      (CHIR ) CHIRON CORP.
XX      XX      Houghton M, Choo Q, Kuo G;
XX      XX      WPI; 2002-040268/05.
XX      XX      P-PSDB; AAE22049.
XX      XX      Combination of hepatitis C viral (HCV) antigens, useful in improved
XX      XX      immunoassay for detecting HCV antibodies.
XX      XX      Example 1; Fig 1; 58pp; English.
XX      XX      The invention relates to combination of hepatitis C viral (HCV) antigens
XX      XX      that have a broader range of immunological reactivity than any single HCV
XX      XX      antigen. The combinations consist of an antigen from the C domain of the
XX      XX      HCV polypeptide, and at least one additional HCV antigen from either the
XX      XX      NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
XX      XX      the form of fusion protein, a simple physical mixture, or the individual
XX      XX      antigens commonly bound to a solid matrix. The combinations of antigens
XX      XX      provides broad range immunoassays for anti-HCV antibodies. The invention
XX      XX      therefore provides a method for detecting antibodies to HCV in a mammal
XX      XX      suspected of containing such antibodies. The present sequence is HCV
XX      XX      polypeptide cDNA
XX      XX      Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;
SQ
XX
XX      Alignment Scores:
XX      Pred. No.: 1,936-241 Length: 9401
XX      Score: 3574.00 Matches: 672
XX      Percent Similarity: 99.42% Conservative: 10
XX      Best Local Similarity: 97.96% Mismatches: 4
XX      Query Match: 98.78% Indels: 0
XX      DB: 6 Gaps: 0
XX
XX      US-09-930-591-2 (1-686) x AAD35043 (1-9401)
Qy      1  MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3417  CTGGCGCCATCAACGGGTACCGCCAGCAAGAGGGGCTCTTAAGGTGCATATACACC 3476
Qy      21  SerIeuThrGlyAlaArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      3477  AGCCTAACTGGCCGGGCAAAAACCAAGTGAAGGGTACGCTCCAGATTGTGTAACCTGCT 3536
Qy      41  AlAGlnThrPheIleuIleAlaThrCysIleAsnGlyValCysThrPheValTyrHisGlyAla 60
Db      3537  GCCCAAACTTCTCTGGCAAGTGCATCAATAGGGGTGGCTGGACGTGCTACACGGGGCC 3596
Qy      61  G1YThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      3597  GGAACGAGACCAATCGGTACCCAGAGGCTCTGTATCAAGTATACCAATGTATGAC 3656
Qy      81  G1AspIleuValGlyTyrProAlaProGlnGlyAlaArgSerIeuThrProCysThrCys 100
Db      3657  CAAGACCTTGTGGGTGGCCGCTCCGCAAGGTAGACCGCTCATTTGACACCCCTCACCTTGC 3716
Qy      101  G1YSerSerIeuThrIleuValThrArgHisAlaAspValIleProValArgArgArg 120

```


Db 3717 GGCCTTCGGAACCTTACCTGGTCAAGAGCAGCCGATGTCATTCCTCCGCGCGCGG 3776
 Qy 121 G1YAspG1YArG1YSerLeuSerProArProIleSerTYrLeuYsG1YSerSer 140
 Db 3777 GGATGATGACAGGGGACGCTGCTGTCCGCCGCCATTTCTTACCTTGAAGAGCTCTCG 3836
 Qy 141 G1YG1YProLeuLeuCySProAlaG1YH1a1aValG1Y1IePheArG1a1a1aValCys 160
 Db 3837 GGGGGTCCGCTGTGTGCTCCCGCGGGGACGCGCTGGGCAATTTAAGGGCGCGCTGTGC 3896
 Qy 161 ThrArG1YVal1a1a1YsAlaValAspPhe1IeProValG1uSerLeuG1uThrThMet 180
 Db 3897 ACCGGTGAAGTGGCTAAAGCGGTGACTTATTCCTGTGAAGAACTTAAGACACCATG 3956
 Qy 181 ArgSerProVal1PheSerAspAsnSerSerProProAlaVal1ProG1nSerTYrG1nVal 200
 Db 3957 AGGCCCCCGGTGTTACAGGATTAATCTCTTCCACCAATAGTCCCGAGCTTCCAGTGG 4016
 Qy 201 AlaH1sleuH1sAlaProThrG1YSerG1YUysSerThrLYsVal1ProAla1a1TYrAla 220
 Db 4017 GCTCACTCCATGCTCCACAGCGAGCGGCAAAAGCAACCAAGTCCCGGCTGCATATGCA 4076
 Qy 221 AlaG1nG1YrLYsVal1leuVal1leuAsnProSerVal1a1a1aThrMetG1YPhG1Y 240
 Db 4077 GCTCAGGCTATAGAGTGTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGTG 4136
 Qy 241 AlaTYrMetSerLYsAlaH1sG1Y1IeAspProAsn1IeArG1ThrG1YVal1ArG1Thr1Ie 260
 Db 4137 GCTTACATGTCACAGGCTCATGGATGCAATCTTACATCAGACCGGGGTGAGAACAAAT 4196
 Qy 261 ThrThrG1YSerProIleThrTYrSerThrTYrG1YLYsPheLeuAlaAspG1YG1YCys 280
 Db 4197 ACCACGTGGACGCCCATCACTACCTCAACCTCAAGCAAGTTCTTGGCGAGCGGGGTGC 4256
 Qy 281 SerG1YG1YAlaTYrAsp1Ie1Ie1IeCysAspG1YUysH1sSerThrAspAlaThrSer 300
 Db 4257 TCGGGGGGCGCTTATGACATATATTGTGACAGAGTGCACCTCCAGGATGCCACATCC 4316
 Qy 301 1IeLeuG1Y1IeG1YThrVal1leuVal1a1aG1uThrAlaG1YAlaArG1LeuThrVal 320
 Db 4317 ATCTTGGGACATGGGACCTGCTCTTGAACCAAGACAGACTGGGGGGGAGACTGGTTGTG 4376
 Qy 321 LeuAlaThrAlaThrProProG1YSerVal1ThrVal1ProH1sProAsn1IeG1uG1uVal 340
 Db 4377 CTGGCCACCGCCACCCCTCCGGGCTCCGTACCTGTGCCCATCCCAACATCGAGAGATT 4436
 Qy 341 AlaLeuSerThrThrG1YG1u1IeProPheTYrG1YsAla1IeProLeuG1uAla1Ie 360
 Db 4437 GCTCTGTCCACCAACCGAGAGATCCCTTTTTCAGGCAAGGCTATCCCTCGAAGTATC 4496
 Qy 361 LysG1YG1YArG1H1sLeu1IePheCysH1sSerLYsLYsCysAspG1uLeuAla1a 380
 Db 4497 AAGGGGGGAGACATCTCATCTTCTGTCAATTCANAGAGAGTGCAGAGACTGGCGCA 4556
 Qy 381 LysLeuVal1a1aLeuG1YVal1a1aVal1a1aTYrTYrArG1YLeuAspVal1SerVal 400
 Db 4557 AAGCTGTCGATTTGGGACATAAGCCGTGSCCTACACCGCGCTCTTGAAGTTCGCTC 4616
 Qy 401 1IeProThrSerG1YAspVal1Val1Val1a1aThrAspAlaLeuMetThrG1YPhThr 420
 Db 4617 ATCCCGACCAAGCGCGATGTTGTGTGTGCAACCATGCTCCCTCAAGACCGGCTATACC 4676
 Qy 421 G1YAspPheAspSerVal1IeAspCysAsnThrCysVal1ThrG1uThrVal1AspPheSer 440
 Db 4677 GGGCACTTCGACTGGGATGACATGACATGCTGTGTGCACCCAGACAGTGCATTTACGC 4736
 Qy 441 LeuAspProThrPheThr1IeG1uThr1IeThrLeuProG1nAspAlaVal1SerArG1Thr 460
 Db 4737 CTTGACCTTACCTTCAACATTAAGACATACGCTCCCGCAGAGTGTGTCTCCGCACT 4796
 Qy 461 G1nArG1ArG1YArG1ThrG1YArG1YLYsProG1Y1IeTYrArG1PheVal1a1ProG1Y 480
 Db 4797 CAAAGTCGGGGCAGAGACTGGCAGGGGGAAGCCAGGCACTTACAGATTGTGTGCACCGGG 4856

Qy 481 G1nArProSerG1YMetPheAspSerSerVal1leuCysG1uCYsTYrAspAlaG1YCys 500
 Db 4857 GAGCGCCCTCCCGGATGTTCCAGCTGCTGTCGTGTGAAGTATGACGAGGCTGT 4916
 Qy 501 AlaTYrTYrG1uLeuThrProAlaG1uThrThrVal1ArG1uThrArG1YrMetLeuThr 520
 Db 4917 GCTTGTATGATGATCCAGCGCCCGGAGACTTACGTTAAGCTTACAGAGTACGAACACC 4976
 Qy 521 ProG1YLeuProVal1CysG1nAspH1sLeuG1uPheThrProG1YVal1PheThrG1YLeu 540
 Db 4977 CCGGGCTTCCCTGTGTCAGAGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCTTC 5036
 Qy 541 ThrH1s1IeAspAlaH1sPheLeuSerG1nThrLYsG1nSerG1Yg1uLeuLeuProTYr 560
 Db 5037 ACTCATATGATGATCCCACTTATCCAGACAAACAGAGTGGGAGAACTTCTTAC 5096
 Qy 561 LeuVal1a1aTYrG1nAlaThrVal1CYsAlaArG1AlaG1nAlaProProProSerTPAsp 580
 Db 5097 CTGGTAGCGTACCAAGCCACCGGTGCGCTAAGGGCTCAAGCCCTCCCATGCTGGAGC 5156
 Qy 581 G1nMetTPAspCysLeu1IeArG1LeuLYsProThrThrLeuH1sG1YProThrProLeuLeu 600
 Db 5157 CAGATGTGAAGTGTGATTGCTCCCTCAAGCCCATGCGTACAGACCCCTGCTA 5216
 Qy 601 TYrArG1uG1YAlaValG1nAsnG1uVal1ThrLeuThrH1sProVal1ThrLYsTYr1Ie 620
 Db 5217 TACAGACTGGGCGCTGTGAGATGAATATCACTTACCGCACCGCAATATCATC 5276
 Qy 621 MetThrCysMetSerAlaAspLeuG1uVal1ThrSerThrThrTPVal1leuVal1G1YArG 640
 Db 5277 ATGACATGCAATGTGCGCGCAGCTGAGGTGCTGACAGACACCTGGGTGCTGTGGCGGC 5336
 Qy 641 ValLeuAla1a1aLeuAla1a1aTYrCysLeuSerThrG1YCyAspVal1a1IeVal1G1YArG 660
 Db 5337 GTCTGCTGCTTGTGGCGCGATGCTGCTGTCAACAGGCTGCGGTGATAGTGGGAGG 5396
 Qy 661 1IeVal1leuSerG1YLYsProAla1Ie1IeProAspArG1uVal1leuTYrArG1uPhe 680
 Db 5397 GTGCTCTGTGCGGGAAGCGGCAATCATATCCGACAGGGAATCCCTTACGAGAGTTCC 5456
 Qy 681 AspG1uMetG1uG1uCYs 686
 Db 5457 GATGAGATGGAAGATGTC 5474

RESULT 15
 AAN92103 standard; DNA; 6905 BP.
 ID AAN92103
 AC AAN92103;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1990 (first entry)
 XX
 DE Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
 DE clones 12f through 15e.
 XX
 KW Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH.
 OS Hepatitis C virus.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 3..6905
 FT /*tag= a
 XX
 PN EP318216-A.
 XX
 PD 31-MAY-1989.
 XX
 PF 18-NOV-1988; 88BP-00310922.
 XX
 PR 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.

PR 26-FEB-1988; 88US-00161072.
 PR 06-MAY-1988; 88US-00191263.
 PR 26-OCT-1988; 88US-00263584.
 PR 14-NOV-1988; 88US-00271450.
 XX
 PA (CHIR) CHIRON CORP.
 PA (CHIR) CHIRON CORP.
 XX
 PI Houghton M, Choo QL, Kuo G;
 XX
 DR WPI: 1989-159274/22.
 DR P-PSDB; AAP92047.
 XX
 PT Purified hepatitis C virus - and associated nucleic acids and
 PT polypeptide(s).
 XX
 PS Claim 3; Fig 32-1 - 32-7; 139pp; English.
 XX
 CC It is a double-stranded nucleotide sequence of the open reading frame
 CC (ORF) (tag a) extending through clones 12f to 15e of hepatitis C virus
 CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
 CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)
 CC it encodes could be used as immunoassay reagents and vaccines and to
 CC generate antibodies useful in diagnosis and passive immunotherapy for HCV
 CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC
 XX
 SQ Sequence 6905 BP; 1421 A; 2082 C; 1946 G; 1456 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 2.16e-241 Length: 6905
 Score: 3571.00 Matches: 671
 Percent Similarity: 99.42% Conservative: 11
 Best Local Similarity: 97.81% Mismatches: 4
 Query Match: 98.70% Indels: 0
 Gaps: 0
 DB: 1
 US-09-930-591-2 (1-686) x AAN92103 (1-6905)
 QY 1 MetLAProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1203 CTGGCGCCCATCAGCGCTACGCCGACGACAAAGGGCTCTTCAGGAGTGCATATCACC 1262
 QY 21 SerLeuThrGlyArgAspGlyAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 1263 AGCTTAAGTGGCCGGGCAAAACCAAGTGAAGGTGAGTCCAGATTGTGTCACTGCT 1322
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleYrHisGlyAla 60
 DB 1323 GCCCAAACTTCCTGGCAACCTGCATCATATGGGTGTGCTGACTGTCTACCAAGGGGCC 1382
 QY 61 GlyThrArgThrIleAlaSerProGlySerProValIleGlnMetYrThrAspValAsp 80
 DB 1383 GGAACGAGACCATGCGCTCACCAAGGGTCTGTCTCATCCAGATGATACCAATGTAGAC 1442
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 1443 CAAGACCTTGGGCTGGCCGGCTCCGCAAGTAGCCGCTCATTTGACACCTGCACCTTGC 1502
 QY 101 GlySerSerAspLeuYrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 1503 GGGCTCTCGGACCTTTACTGTGTACGCGACGACGCCCATGTCTATCCGTGCGCGGGCG 1562
 QY 121 GlnAspGlyArgGlySerLeuLeuSerProArgProIleSerYrLeuLeuGlySerSer 140
 DB 1563 GGTATATGACAGGGGCGAGCTGTCTGTCTGTCCCGCCCATTTCTATCTTGAAGAGCTCTCG 1622
 QY 141 GlnGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 1623 GGGGGTCCGCTGTGTGTGCGCGCGGGGACGCCCTGGGCAATTTAAGGCGCGGTGTGC 1682
 QY 161 ThrArgGlyValAlaValAspPheIleProValGluSerLeuGluThrThrMet 180

DB 1683 ACCCGTAGAGTGGCTAAGCGGTGACCTTATCCCTGTGGAGAACCTAGACACCATG 1742
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal 200
 DB 1743 AGGTCGCCGAGTGTACGGATTAATCTCTTCTCCACGATGATGCCCCAGACTTCCAGGTG 1802
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyYrSerThrYrValProAlaAlaYrAla 220
 DB 1803 GCTCACTCATGCTCCCAAGGACAGCGGCAAAAGACCAAGGATCCCGGTGCATATGCA 1862
 QY 221 AlaGlnGlyYrYrValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 1863 GCTCAGGGCTATAAGGTGCTAGTACCAACCCCTGTGTCTGCAACACTGGGCTTTGGT 1922
 QY 241 AlaTyrMetSerYrValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 1923 GCTTACATGTCCAAAGCTCATGGATTCGATCTTAACTCAGACCGGGGTGAACAATT 1982
 QY 261 ThrThrGlySerProIleThrYrSerThrYrGlyYrPheLeuAlaAspGlyCys 280
 DB 1983 ACCACTGGACGCCCATCAGTACCTCACTCAGGCAAGTTCCTTCCGACGGCGGTGC 2042
 QY 281 SerGlyGlyAlaYrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 2043 TCGGGGGGCGCTTATGACATATATTTGTGACACTGCACCTCCACGATGCCATATCC 2102
 QY 301 IleLeuGlyYrIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
 DB 2103 ATCTTGGGATTCGGACATGTCTCTTGAACAGACGACCTGGGGGCGAGACTGTGTG 2162
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnVal 340
 DB 2163 CTGCGCACCGCCACCCCTCGGGCTCCGTACACTGTGCCCATCCCAATCCAGAGGTT 2222
 QY 341 AlaLeuSerThrThrGlyGluIleProPheYrGlyYrAlaIleProLeuGluAlaIle 360
 DB 2223 GCTCTGTCCACCAACCGAGAGATCCCTTTTAAAGGACGATATCCCTCGAAGTAAATC 2282
 QY 361 LysGlyGlyYrArgHisLeuIlePheCysHisSerYrLysYrCysAspGluLeuAlaAla 380
 DB 2283 AAGGGGGGAGACATCTCATCTTCTGTCTCAAAAGAGGCGACCAATCCGCCGCA 2342
 QY 381 LysLeuValAlaLeuGlyValaAsnAlaValaIleYrYrArgGlyLeuAspValSerVal 400
 DB 2343 AAGCTGGTGCATTGGGCATCATGCGGTGACCTTCAACGCGGTCTTGAACGATCCGTC 2402
 QY 401 IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
 DB 2403 ATCCGACCAAGCGGCGATGTGTGTCGTGGCAACGATGCCCTCATGACCGGCTATACC 2462
 QY 421 GlnAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 2463 GCGGACTTGCATCTCGGTGATGACTGCATACGTGTGTACCCAGACAGTCAATTTCAAG 2522
 QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 2523 CTTCACCCATCTTCAACATATGACATCAACGCTCCCAAGATCTGTCTCCGACACT 2582
 QY 461 GlnArgArgGlyArgThrGlyArgGlyYrProGlyYrIleYrArgPheValAlaProGly 480
 DB 2583 CAACGTGCGGGGAGGACCTGGCAGGGGAAAGCCAGGATCTTACGATTTGTGGACCGGGG 2642
 QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGlyCysYrYrAspAlaGlyCys 500
 DB 2643 GAGCGCCCTCCGCGCATGTTCACATCTGTCTGTGTGAGTGTCTATACGAGGCTGT 2702
 QY 501 AlaTrpYrGluLeuThrProAlaGluThrThrValArgLeuArgAlaYrMetAsnThr 520
 DB 2703 GCTTGGATAGCTCAAGCCCGCCAGACCTACGTTAAGCTACGACGTAACGAACACCC 2762
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 DB 2763 CCGGGGCTTCCCGTGTGCAAGACCATCTTGAATTTTGGAGGGCGTCTTTTACAGGCTTC 2822

```
Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyIleAsnLeuProTyr 560
      |||
Db      2823 ACTCATATAGATGCCCACTTCTATATCCAGCAAGAGAGTGGGAGAACTTCCCTTAC 2882
      |||
Qy      561 LeuValAlaIleYrgIleAlaThrValCysAlaArgAlaGlnAlaProProPheSerTTPAsp 580
      |||
Db      2883 CTGGTAGCCGTAACCAAGCCAGCGTGTGCTAGGGCTCAAGCCCCCTCCCATCGTGGAC 2942
      |||
Qy      581 GlnMetTrpLysCysLeuIleArgIleuLysProThrIleuHisGlyProThrProLeuLeu 600
      |||
Db      2943 CAGATGTGAGAGTGTGATTGATTGCTCAGGCCACCTCCATGGGCCAACACCCCTGCTA 3002
      |||
Qy      601 TyrArgIleuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIleTyrIle 620
      |||
Db      3003 TACAGACTGGGCGCTGTTCAGAAATGAATCACCCCTGACGCAACCAATATCATC 3062
      |||
Qy      621 MetThrCysMetSerAlaAspLeuGluValIleThrSerThrTyrValLeuValGlyIle 640
      |||
Db      3063 ATGACATGCATGTCTGGCCGACCTGGAGGTGCTCACGACGACCTGGGTGCTGTTGGCGGC 3122
      |||
Qy      641 ValLeuAlaAlaLeuAlaAlaIleTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
      |||
Db      3123 GTCTGGCTGCTTGGCGCGATATGCTGTCAACAGGCTGCGTGTATAGTGGCGAGG 3182
      |||
Qy      661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
      |||
Db      3183 GTCGTCTGTCCGGGAAGCCGGCAATCATCTGTACAGGGAAGTCTTACCGAGAGTTTC 3242
      |||
Qy      681 AspGluMetGluGluCys 686
      |||
Db      3243 GATGAGATGGAGAGTGC 3260
      |||
```

Search completed: September 17, 2004, 08:29:52
Job time : 766 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 07:46:20 ; Search time 127 Seconds
(without alignments)
2997.609 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPITAYAOCTRLGLCIIIT.....PAIIPPREVLYREFDEMEEC 686

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO_epool/US09930591/rnat 13092004_164953_711/app query.fasta_1.839
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rni -WINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum2 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09930591 @CGN 1.1 56 @rnat 13092004_164953_711 -NCP=6 -ICPU=3
-NO_MAMP -IARSEQOTERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGCLOG
-DEV_TIMEOUT=120 -WASN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3574	98.8	7310	3	US-08-444-818-74
2	3574	98.8	9379	3	US-09-388-874-1
3	3574	98.8	9379	4	US-09-916-359-1
4	3574	98.8	9401	1	US-07-910-760-9
5	3574	98.8	9401	1	US-08-440-519-9
6	3574	98.8	9401	4	US-08-440-549-9
7	3574	98.8	9401	4	US-08-823-895A-25
8	3571	98.7	9401	4	US-08-444-818-65
9	3570	98.7	8316	3	US-08-444-818-88
10	3570	98.7	9185	3	US-08-444-818-122
11	3570	98.7	9185	3	US-08-444-818-123
12	3565	98.5	2058	4	US-09-881-239-2

13	3565	98.5	2058	4	US-09-881-654-1
14	3565	98.5	8967	3	US-08-444-818-137
15	3555	98.3	5360	3	US-08-444-818-53
16	3551	98.1	9646	4	US-08-811-566-1
17	3551	98.1	9646	4	US-09-034-756-1
18	3551	98.1	12980	3	US-08-811-566-5
19	3551	98.1	12980	3	US-09-034-756-5
20	3544	98.0	9599	3	US-09-014-416-2
21	3544	98.0	9599	3	US-09-014-416-6
22	3538	97.8	9401	2	US-08-432-693-1
23	3538	97.8	9416	3	US-08-811-566-19
24	3538	97.8	9416	4	US-09-034-756-19
25	3535	97.7	9379	3	US-08-444-818-176
26	3532	97.6	9401	5	PCT-US91-02225-9
27	3504	96.8	9416	4	US-08-823-895A-26
28	3504	96.8	9416	4	US-10-104-966-13
29	3429	94.8	7989	4	US-09-539-601-10
30	3429	94.8	8001	4	US-09-539-601-7
31	3426	94.7	8637	4	US-09-539-601-4
32	3426	94.7	8649	4	US-09-539-601-13
33	3426	94.6	11076	4	US-09-539-601-1
34	3423	94.6	8001	4	US-09-539-601-22
35	3420	94.5	6039	1	US-08-324-977-11
36	3420	94.5	6039	2	US-08-384-616-11
37	3420	94.5	6039	2	US-08-904-686A-11
38	3420	94.5	6039	3	US-09-315-850-11
39	3420	94.5	7863	1	US-08-324-977-35
40	3420	94.5	7863	2	US-08-384-616-35
41	3420	94.5	7863	2	US-08-904-686A-35
42	3420	94.5	7863	3	US-09-315-850-35
43	3420	94.5	7917	1	US-08-324-977-31
44	3420	94.5	7917	2	US-08-384-616-31
45	3420	94.5	7917	2	US-08-904-686A-31

ALIGNMENTS

RESULT 1
US-08-444-818-74
; Sequence 74, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; INVENTOR: Rutter, William J.
; TITLE OF INVENTION: NANBY Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chilton Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Aisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:
LENGTH: 7310 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..7310
US-08-444-818-74

Alignment Scores:

Pred. No.:	0	Length:	7310
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	3	Gaps:	0

US-09-930-591-2 (1-686) x US-08-444-818-74 (1-7310)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
DB ::
1728 CTGGCGCCCATCAGCGGTACGCCAGACAGAGGGCCCTCTAGGGTGCATATACAC 1787
QY 21 SerLeuThrGlyATGAspLysAsnGlnValGlnGlyGluValGlnIleValSerThrAla 40
DB AGCTTAATGCGCCCGGAGCAAAACCAAGTGAAGGGTGAAGTCCAGATTGTGTCAACGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB GCCCAAACTTCTCGGCAACGTCATCATGGGGTGTGTGAGCTGTCTACACAGGGGCC 1907
QY 61 GlyThrArgTrpTrpIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB GGAACGAGACCATCGCTCACCAGAGGTCCTGTCTCATCCAGATGATACCAATGATGAC 1967
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB CAACACCTTGTGGGCTGGCCCGCTCCGAGAGTACCGGCTCATTTAGACCCCTGACATTGC 2027
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB GGCTCCCTCGGACCTTACCTGTGTACAGAGGACCGCCATGTCTCCCGTGGCCGGCGG 2087
QY 121 GlyAspArgArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB GGTAATACAGAGGGGAGGCTGTGTGTGCGCCGCGCCATTCTCTAATTGAAGGGCTCTCG 2147
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB GGGGGTCCGCTGTTGTGTGCCCCGGGGGACGCGTGGGCATATTAGGGCGCGGTGTGC 2207
QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGlnSerLeuGluThrMet 180
DB ACCCGTGAAGTGGCTAAGCGGTGACTTATTCCTGTGAGAACTTAGAGCAACCATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB AGGTCCCGGCTGTTCAGGATTAACCTCTCTCACAGATGAGCCAGAGCTTCCAGAGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB GCTACCTCCATGTCTCCACAGGAGGAGGCAAAAGACCAAGTCCCGGCTGATAGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB GCTAGAGGCTATTAAGTGTAGTACTCAACCCCTCTGTGTGCAAACTAGGGCTTTGGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyTyrLeuAspProsnIleArgThrGlyValArgThrIle 260
DB GCTTAACATGTCCAAAGGCTCATGGATCGATCTTAACATCAGACCGGGGTGAGAACAATT 2507

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB ACCACTGGACAGCCCCCATCAAGTACTCACTCACGGAAGTTCCTTCCGACGGCGGGTGC 2567
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB TCGGGGGGGCTTATACATATATATTGTGACGAGTGTCCACTCCAGATGACACATCC 2627
QY 301 IleLeuGlyTyrIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB ATCTTGGGCATGGGACGTGTCTCTTGACCAAGACAGAACTCGGGGGCGAGACTGTGTG 2687
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB CTGCGCACCGCCACCCCTCGGGGCTCGTCACTGTGCCCATCCCAACTCCAGAGAGTT 2747
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB GCTGTGTCCACCCAGAGATCCCTTTTACGGGACGACTATCCCTCCGAGTAAATC 2807
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysGlyLysCysAspGluLeuAlaAla 380
DB AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAAGATGCGACGAACTCGCGCA 2867
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB AACTGGTGGCATTTGGGACATGACCGGTGGCTTACCGCGGCTTGTACGTGTCCGTC 2927
QY 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB ATCCACACAGCGCGCATGTGTGTGTGGCAACGAGTCCCTCATGACCGGCTATATCC 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB GCGGACTTGCACTCGGTGATGACTGCATACGTGTGTACCCAGACATGCTTACAG 3047
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB CTTGACCTTACCTTCAACATTGAGACATACACGCTCCCGAGATCTGTCTCCGCACT 3107
QY 461 GlnAspArgGlyArgTrpGlyArgGlyLysProGlyTyrLeuArgPheValAlaProGly 480
DB CAACGTGGGGGAGACTGAGCGAGGGGAAAGCCAGGATCTTACAGATTGTGGACGGGG 3167
QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyLysCysTyrAspAlaGlyCys 500
DB GAGCGCCCTCCCGGCATGTTTGCATCTGTCTGTGTGAGTGCATGACGAGGCTGT 3227
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB GCTTGGTATGAGCTCACGCGCCGAGACTACAGTTAAGGCTACGAGCGTACATGAACCC 3287
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGlnGlyValPheThrGlyLeu 540
DB CCGGGGCTTCCGTGTGCGAGGACATCTTAATTTTGGAGGGCTTTTACAGGCTTC 3347
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB ACTCATATAGATGCCCACTTCTATCCAGACAAACAGAGTGGGAGAACTTCTTTC 3407
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB CTGGTAGCGTACCAACCGGCTGCTAGGGGCTCAAGCCCTCCCATCTGTGGGAGC 3467
QY 581 GlnMetTrpLysGlyLeuIleAspGlyLeuSerProThrLeuHisGlyProThrProLeuLeu 600
DB CAGATGTGAAGTGTGATTCCTCAAGCCCACTTCATGGGCAACCCCTGCGTA 3527
QY 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisIleProValThrLysTyrIle 620
DB TACAGACTGGGCGCTGTGTCAAGATGAATACACCTTACGACCCATCCAAATACATC 3587
QY 621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTrpValLeuValGlyGly 640

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCyGluCysTyrAspAlaGlyCys 500
DB 4835 GAAGGCCCTCCCGCATGTTGAGCTGCTGCTCTGAGAGGATGATGACGAGGCTCT 4894
QY 501 AlaTyrPylGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4895 GCTTGATAGAGCTCAAGCCCGGAGAGACTAGATTAGCTACGAGAGCTACGATGAAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrPylGluValPheThrGlyLeu 540
DB 4955 CCGGGGCTTCCTCGTTCGACGACCACTTTGGAATTTGGGAGGCGCTTTACAGGCTTC 5014
QY 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrThrGlnSerGlyGluAsnLeuProTyr 560
DB 5015 ACTCATATGATGCGCACTTCTATCCAGACAAAGCAGAGTGAGGAGAACTTCTCTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrPyl 580
DB 5075 CTGGTACCGTACCAAGCCAGCTGTGCGCTAGGCTCAAGCCCTCCCATGCTGGGAC 5134
QY 581 GlnMetTyrPylCysLeuLeuArgLeuLeuPheThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGCGAAGCTTTGATTGCTGCTCAAGCCCACTCATGGGCGAAGACCCCTGCTA 5194
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrTyrTyrIle 620
DB 5195 TACAGACTGGGCGCTGTTGAGATGAAATACCTCGAGCACCAGTCACCAATATCATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGlyValValThrSerThrTyrValLeuValGlyGly 640
DB 5255 ATGACATCATGTGCGCGGACCTGGAAGTCTGTCACGAGCACCTGGGCTGCTGGCGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 5315 GTCCGCTGCTTTGGCCGCGATGCTGTCACAGGCTCGGTCATATGAGGACGG 5374
QY 661 IleValIleuSerGlyLeuProAlaIleThrProAspArgGluValLeuTyrArgGluPhe 680
DB 5375 GTGCTGCTGTCGGGAGGCGGCAATCATCTGACAGGGAAGTCTCTACGAGAGTTC 5434
QY 681 AspGluMetGluGluCys 686
DB 5435 GATGAGATGGAAGAGTGC 5452
RESULT 3
US-09-916-359-1
; Sequence 1, Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; FILE REFERENCE: PWC97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320)...(9352)
US-09-916-359-1
Alignment Scores: 0 Length: 9379
Pred. No.:

Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 4 Gaps: 0
US-09-930-591-2 (1-686) x US-09-916-359-1 (1-9379)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3395 CTGGCGCCCATCAGCGCTGATGAGCCAGCAGCAAGGGGCTCTTACGGGTGATATACACC 3454
QY 21 SerLeuThrGlyArgAspPylAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 3455 AGCTTAACCTGGCGGAGCAAAACCAAGTGAAGGTGAGTCCAGATGTGTCACTGCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
DB 3515 GCCCAACCTTCTGGCAACCTGCACTCAATGGGTGTGCTGAGCTGTACACGAGGGCC 3574
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3575 GGAAGAGAGCATGCGCTCAACCAAGGTCCTGTCAATCCAGATGTATACCAATGTAGAC 3634
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3635 CAAGACCTTGTGGCTGGCTGCTCGCAAGTACCGCTCATTTGACACCTGCACTTGC 3694
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 3695 GGCTCTCGGACTTTACTGCTGTCACAGGACCGCGATGTATCCGTGGCGCGCGG 3754
QY 121 GlyAspGlyArgGlySerLeuLeuSerProAlaGlyProIleSerTyrLeuLeuGlySerSer 140
DB 3755 GGTGATAGCAGGGGCGCTGCTGTCGCGCGCCCATTTCTTACTTGAAAGCTCTCG 3814
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3815 GGGGCTCGCTGTGTGTCGCCCGGGGACGCGGTGGCATTTAAGGCCCGCGGTGTC 3874
QY 161 ThrArgGlyValAlaAlaValAlaAspPheIleProValGluSerLeuGlnThrThrMet 180
DB 3875 ACCCGTGAAGTGGCTAAGCGGCGGTGACTTATCCCTGTGAAGAACCTTACGACACATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3935 AGGTCCCGGTGTTCAAGGATATCTCTCTCCACGATGTGCTCCAGAGCTTCCAGGTG 3994
QY 201 AlaHisPheHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3995 GCTACCTCCATGCTCCCAAGCAGCGGCAAAAGCAGCAAGGTCCCGGTGATATGCA 4054
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4055 GCTCAGGGCTATAGGAGGTGATGACTCAACCTCTGTGTCTCAACACGTGGCTTTGTG 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgTyrIle 260
DB 4115 GCTTACATGTCAGGCTCATGGATGATGATCTTAACATCAGACCGGGGTGAGAAACATTT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 4175 ACCACTGGACCCCATCAGTACCTCCACTACGCGCAAGGTCTTGGCGAGGGGTGC 4234
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 4235 TCGGGGGCGCTTATGACATATATATTTGACAGAGTGCACCTCCAGGATGACATCC 4294
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4295 ATCTTGGGCACTGGCACTGTCTTGCACCAAGCAGAGACTCGGGGGGAGACCTGGTGTG 4354
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340

Db 4355 CTGGCCAGCCGACCCCTCGGCGTCGTCATGTCGCCATCCCAATCCGAGGAGTT 4414
Qy 341 AAlaSerThrThrGlyGluIleProPheTyrGlyValAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTGTCCACCAACGAGAGATCCCTTTTACGGCAAGCTATCCCCCTCGAAGTATC 4474
Qy 361 LysGlyValArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaIle 380
Db 4475 AAGGGGGGAGACATCTCATCTTCTGTGATTAAGAGAAAGTGGCGAATCTGGCGCA 4534
Qy 381 LysLeuValAlaLeuGlyValAlaSerAlaAlaIleTyrThrArgGlyLeuAspValSerVal 400
Db 4535 AAGCTGTCTCCATTTGGCATCAATGCGGTGGCTACTACGCGGTCTTGAACGTGCGCTC 4594
Qy 401 ILProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCCAACCAAGCGGCAATGTTGCTGCTGGCAACGATGCCCTCATGACCGGCTATACC 4654
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GCGCACTTGACCTCGGTGATGACCTGCAATACGTGTTCACCCAGACAGTCGATTCAGC 4714
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCTTACCTTCAACATTTAGACATACGCTCCCGAGATGCTGCTCCGCGACT 4774
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4775 CAACGTCGGGGGAGACCTGGCAAGGGGAGAACCGGACCTTACAGATTTGGGCAACGGGG 4834
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4835 GACGCGCCCTCCGCAATGTTGCACTGCTCCCTCTGTAGAGCTATGACGCAAGGCTGT 4894
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4895 GCTTGGATAGCTCAACGCGCGGAGACATCACTTACGAGCGTACGAGCGTACATGAACAC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGGCTTCCCGTGGCCAGACCACTTGAATTTGGAGAGCGCTTTTACAGGCGTTC 5014
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATATATATCCCACTTCTATCCCAAGAGAGAGTGGGAGAACCTTCTTAC 5074
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrPasp 580
Db 5075 CTGGTAGCGTACCAAGCCACCGTGTGGCTAGGAGCTCAAGCCCTCCCATCGTGGGAC 5134
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAAGGTTGATTTGCTTCACCTCAAGCCCACTTCATGGGCAACCCCTGCTA 5194
Qy 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGGCGCTGTTCAAGATGAATCACTTCAGACCACTCAGTCAACCAATACATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 5255 ATGACATGACATGTCGCGCGACCTGAGGTGCTGACAGGACCTGGGCTCGTGGGCGG 5314
Qy 641 ValLeuAlaAlaLeuAlaAlaIleTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGTGCTGCTTTGGCGGGTATGCTGTTCACAGGCTGCGTGTGCATATAGGGCAGG 5374
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTGCTCTGTCTCGGAAAGCCGCAATCATCTGACAGGGAAGTCTCTACCAAGATTC 5434
Qy 681 AspGluMetGluGluCys 686

Db 5435 GATGAGATGAGAGATGTC 5452
RESULT 4
US-07-910-760-9
; Sequence 9, Application US/07910760
; Patent No. 5683864
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/910,760
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Eq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 342..9374
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366-
; OTHER INFORMATION: /note="This amino acid position
; OTHER INFORMATION: can also be Arg."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 372-
; OTHER INFORMATION: /note="This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 867
; OTHER INFORMATION: /note="This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1341
; OTHER INFORMATION: /note="This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2148
; OTHER INFORMATION: /note="This amino acid position
; OTHER INFORMATION: can also be Ile."
; FEATURE:
; NAME/KEY: misc_feature

```

LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."

```

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-07-910-760-9

Alignment Scores:
Pred. No.: 0
Score: 3574.00
Percent Similarity: 99.42%
Best Local Similarity: 97.96%
Query Match: 98.78%
DB: 1
Gaps: 0

US-09-930-591-2 (1-686) x US-07-910-760-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3417 CTGGCCCATCATCCGCGTCACCCGACGACGACGAGGGGCTCTCGAGGTGCATATCACC 3476
QY 21 SerLeuThrGlyArgAspIysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 3477 AGCCTAATCGCCGCGGCAAAACCAAGTGGAGGTGAGGTCAATATGTGCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 3537 GCCCAACCTCTCTGGCAACGTGCATCATATGGGTGTCTGACGTGTACACGAGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3597 GGAACGAGGACATCGTCGACCAAGGGTCTGTATCATCAATATGATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaPProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3657 CAAGACCTGTGTGGCTGGCCGCTCCGCAAGTAGCCGCTCATGTACACCTGCACCTTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAlaArgArg 120
DB 3717 GGCTCTCGGACCTTACCTTACCTGTACAGACGACGCGAATGTCTCCGTGCGCCGCG 3776
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuIysGlySerSer 140
DB 3777 GGTATAGCAGAGGGGAGCTGTGTGCGCCCGGCAATTTCTACTTGAAGGCTCTCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheAsnArgAlaAlaVal 160
DB 3837 GGGGCTCGGCTGTGTGCTCCCGCGGGGACGACCGCTGGCATATTGAGGCGCGGCTCTCG 3896
QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGluThrMet 180
DB 3897 ACCGTGAGAGGTGTAAGCGGTGACCTTATCTCTGTGAGACACTTGAACACCATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3957 AGTCCCGGCTGTACAGATACCTCTTCCACCAATAGTGTCCCGGACCTTCCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlyIysSerThrIysValProAlaAlaTyrAla 220
DB 4017 GCTACCTCCATGCTCTCCACGAGGCGGAAAGACCAAGGTCCGCGCTGCATATGCA 4076
QY 221 AlaGlnGlyTyrIysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

```

Db 4077 GCTCAGGGCTAATAGGTGCTAGTACCAACCCCTCTGTTGCTGCACACATGGGGCTTTGGT 4136
 Qy 241 AATyMeSerlySalahIsglylleAspProhenilleargThrglyValArgThrIle 260
 Db 4137 GCTTACATGTCAGAGGCTCATGGATGATCTTACATCAGAGCCGGGGTGAAGACATTT 4196
 Qy 261 ThrtHrglySerProilethrtYrSerThrtYrlylySpheluaIaAspglylyCys 280
 Db 4197 ACCACTGGACGCCCATCAGTACTCCACCTAGCGCAAGTTCTTGGCGAGGGGGGTGC 4256
 Qy 281 SerGlylyValArgpillelleleCysAspglylyCysHisSerThraPalaThSer 300
 Db 4257 TCGGGGGGCGCTTATGACATTAATTTGTGACAGAGCCATCTCCAGAGTGCACATCC 4316
 Qy 301 lleleuglylleIleThraValleuAspglylyValarglylyValaarglyleThraVal 320
 Db 4317 ATCTTGGGACATGGGACCTGTCTTGGACCAAGAGACTGGGGGGGAGAGCTGGTTGTG 4376
 Qy 321 leuAlaThraIaThrProProglySerValThraValProHisProAsenilleuglylyVal 340
 Db 4377 CTGGCCACCGCCACCCCTCGGGGCTCCGTACGTGCTCCATCCCAATCGAGAGAGTT 4436
 Qy 341 AlaSerSerThrtHrglylylleProPheThrglylyValalleProleuglylyVal 360
 Db 4437 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTATC 4496
 Qy 361 lyeglylyValrghIleuIlePheCysHisSerlylylySpheluaIaAspglylyVala 380
 Db 4497 AAGGGGGGAGACATCTCATCTTCTGTCTTCAAGAGAGAGTGCAGACATCGCGCA 4556
 Qy 381 lyseuValaIaIeuglylyValaAsnIaValaIaTyrrArglylyleuAspValSerVal 400
 Db 4557 AAGCTGTCCATTTGGGACATCAATGCCGTGGCCCTACTACCGCGCTTTGACGTGTCCGT 4616
 Qy 401 lleProThrSerGlyAspValValaValaIaThraPalaIeueWctThrglyPheThr 420
 Db 4617 ATCCCGACCAAGCGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4676
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrcysValThrglnThraValAspPheSer 440
 Db 4677 GGGACCTTCAGCTCGGATGATGACTGCATACGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4736
 Qy 441 leuAspProThrPheThrIleGluThrIleThrleuProGlnAspAlaIaSerArgThr 460
 Db 4737 CTGACCTTCACCTTCACTTGAAGACATACAGCTCCCGCAGATGTGTGTGTGTGTGTGT 4796
 Qy 461 GlnArgArglylyArgThrglylylyAspProglylylyTyrrArgPheValaIaProgly 480
 Db 4797 CAACGTGGGGGAGAGCTGGCAGGGGAGAGCCAGATCTTACAGATTTGTGTGTGTGTGTGT 4856
 Qy 481 GlnArgProSerGlyMetPheAspSerValleuGlylyCysTyrrAspAlaGlyCys 500
 Db 4857 GAGGGCCCTCCCGGACATGT 4916
 Qy 501 AlaTrpIlyrGluThrProAlaGluThrThraValArgleuArgAlaTyrrMetAsnThr 520
 Db 4917 GCTTGGATAGCTCAGCGCCCGCAGACATACATTTGCTGCTGCTGCTGCTGCTGCTGCTG 4976
 Qy 521 ProGlyleuProValCysGlnAspHisleuGlnPheTrpGlylyValPheThrglylyleu 540
 Db 4977 CCGGGGCTTCCTCGGT 5036
 Qy 541 ThrtHrglyleuPalaHisPheleuSerGlnThrtYrGlnSerGlylyuanleuProTyrr 560
 Db 5037 ACTCATATGATGCTCCATTTTATCCAGCAAGAGAGTGGGGAGAACTTCTTAC 5096
 Qy 561 leuValaIaTyrrGlnaIaThraValCysAlaArgAlaGlnaIaPProProSerTrpAsp 580
 Db 5097 CTGGTACGATACCAAGCAACCGTGTGGCTAGGGCTCAAGCCCTCCCACTGTGGGAC 5156
 Qy 581 GlnMetTrpIlyrCysleuIleArgleuIlyAspProThrleuHisGlyProThrProleu 600
 Db 5157 CAGATGTGAAGT 5216

Qy 601 TyrrArgleuGlyAlaValGlnAsnGlyValThrleuThrHisProValThrtYrIle 620
 Db 5217 TACAGCTGGGGCTCTTCAATGATGAATACACCTTACAGCACCACTGACCAATATCATC 5276
 Qy 621 MetThrCysMeSerAlaAspIleuValValThrSerThrtPalleuValGlyly 640
 Db 5277 ATGACATGATGT 5336
 Qy 641 ValIleuAlaIaIeuglylyValaIaTyrrCysleuSerThrglyCysValIalleValIArg 660
 Db 5337 GTCTGTGCTGT 5396
 Qy 661 lleValleuSerGlylyAspProAlaIlelleProAspArgGlylyleuTyrrArglyly 680
 Db 5397 GTCTGT 5456
 Qy 681 AspglyMetGlylyCys 696
 Db 5457 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5474

RESULT 5
 US-08-440-519-9
 ; Sequence 9, Application US/08440519
 ; Patent No. 5712087
 ; GENERAL INFORMATION:
 ; APPLICANT: Houghton, Michael
 ; APPLICANT: Choo, Qui-Lim
 ; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
 ; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Chiron Corporation
 ; STREET: P.O. Box 8097 (Int. Prop. R-440)
 ; CITY: Emeryville
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94662-8097
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/440,519
 ; FILING DATE: 12-MAY-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/910,760
 ; FILING DATE: 07-JUL-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Blackburn Esq., Robert P.
 ; REGISTRATION NUMBER: 30,447
 ; REFERENCE/DOCKET NUMBER: 0101.002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (510) 601-2702
 ; TELEFAX: (510) 655-3542
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9401 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 342..9374
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 366
 ; OTHER INFORMATION: /note= "This amino acid position
 ; OTHER INFORMATION: can also be Arg."

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5370
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402

```

```

OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-519-9

Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
Gaps: 0
US-09-930-591-2 (1-686) x US-08-440-519-9 (1-9401)
QY 1 MechAlProTlleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGCGCCATCAACGGGCTACGCCCAACAGAGGGCTCTTAAGGTGCATATACAC 3476
QY 21 SerLeuThrgIyArAspIyAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCTTAATGGCCGGGCAAAAACCAAGTGAAGGTGATGTCAGATTGTCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysThrThrValTyrHisGlyAla 60
Db 3537 GCCCAAACTTCTCGGCAACGTGCATCAAAAGGGGTGCTGAGCTGTCTACACGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProIyGlyProValIleGlnIleThrAsnValAsp 80
Db 3597 GGAACGAGACCAATCGCTACCCAAAGGTCCTGTATCCAGATGTATACCAATGTATGAC 3656
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGAGGTGGCCCGCTCCGCAAGTACCGCTCATTGACACCCCTGACTTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAlaArgArg 120
Db 3717 GGCTCTCGAGACTTATCTGTACGTGATCAGAGGACCGCATGTATCCCGTGGCCGGCGG 3776

```

QY 121 GYAASGlyAArglySerLeuLeuSerProAlaProIleSerTyrluSgLySerSer 140
 DB 3777 GGGAGAGACAGGGGAGCCCTGTCTGCTCCCGGCGCATTTCTACTGAAGAGCTCTCG 3836
 QY 141 GYGLProleuLeuYsProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 3837 GGGGGGTCGGTGTGTGCCCCGGGGGACGCCGTGGGCATATTATAGGGCGCGGTGTC 3896
 QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGluThrThmet 180
 DB 3897 ACCCGTAGAGTGGCTAAGGCGGTGAGCTTATTCCTGTGAGAACCTAGAGACACATG 3956
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrluVal 200
 DB 3957 AGGTCCCGGTGTTCACGGATTAACCTCTCTCACAGTAGTCCCGGACCTTCAGGTG 4016
 QY 201 AlaHisLeuHisAlaProThrArglySerGlyLysSerThrlySvalProAlaAlaTyrluAla 220
 DB 4017 GCTCAGCTCCATGCTCCACAGAGCGGCAAAAGCCAAAGTCCCGCTGCATATGCA 4076
 QY 221 AlaGlnGlyTyrluSvalLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 4077 GCTCAGGGCTTAAGGCTAGTACTCAACCCCTGTGTCTGCACACATGGGCTTGGT 4136
 QY 241 AlertyMetSerlySvalHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 4137 GCTTACATGTCGAAGGCTCATGGGATGATCTTAACATCAGAGACCGGGGTGAGAACAT 4196
 QY 261 ThrThrGlySerProIleThrTyrluSerThrTyrluSvalPheLeuAlaAspGlyGlyCys 280
 DB 4197 ACCACTGGACGGCCCATCAGTACTCCACTACGGCAAGTTCCTTGGCAGCGGGGTGC 4256
 QY 281 SerGlyValAlaTyrluAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 4257 TCGGGGGGCGCTTATGACATATATTTGTGACGAGGCTCCACGAGTGCACATCC 4316
 QY 301 IleLeuGlyTyrluGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 4317 ATCTTGGGCGATCGGCACTGCTTGAACAAGAGAGCTCGGGGGGAGACATGGTTGTG 4376
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 4377 CTGGCAGCCGCCACCCCTCGGGCTCGTCACTGGTCCCATCCCAACATCGAGAGCT 4436
 QY 341 AlaLeuSerThrThrArglyGluIleProPheTyrluSvalAlaIleProLeuGluAlaIle 360
 DB 4437 GCTCTGTCCACCCAGAGATCCCTTTTACGAGGCTATCCCTCGAGAGTATC 4496
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerlySvalLysCysAspGluLeuAlaAla 380
 DB 4497 AAGGGGGGAGACATCTCTCTGTCAATAAGAAAGAGTGCAGACGAACTCGCGCA 4556
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrluArgGlyLeuAspValSerVal 400
 DB 4557 AAGCTGTCCTCATTTGGGCATATGCGGTGGCTACTACCGCGTCTTGTGACGTCCGTC 4616
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrArglyPheThr 420
 DB 4617 ATCCGACAGCGCGAGTGTGTGTGTGTGCAACGATGCCCTCATGACCGGCTATAC 4676
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 4677 GCGGACTTGACCTCGGTGATGACTGCAATACGTGTGTGCAACGACAGTGCATTTACG 4736
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 4737 CTTCAGCCCTTACCTTCACTTGAACATCAGCTCCCGGAGATGCTGTCTCCGCACT 4796
 QY 461 GlnArgArgGlyValArgThrArglyArgGlyLysProGlyTyrluThrArgPheValAlaProGly 480
 DB 4797 CAACGTCGGGGAGAGCTGGCAAGGGGAGACCGGAGCATCTACAGATTTGTGGCACCGGG 4856
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuGlyGluCysTyrluAspAlaGlyCys 500

DB 4857 GAGCGCCCTCCCGCATGTTCAGCTGTCCTGCTGTGAGTGTATGACGAGCTGT 4916
 QY 501 AlaTyrluGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrluMetAsnThr 520
 DB 4917 GCTTGGTATGAGCTTACGCGCCCGGAGACTACAGTTAGGCTTACAGACGTAACGAC 4976
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrluGlyValPheThrGlyLeu 540
 DB 4977 CCGGGGCTCCCGGTGTGCCAGAGCATCTTGAATTTTGGAGGGCGCTCTTACAGCGCTC 5036
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysAsnLeuProTyrlu 560
 DB 5037 ACTCATATGATGCCCCACTTTCATCCACAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5096
 QY 561 LeuValAlaTyrluGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrlu 580
 DB 5097 CTGGTAGCTTACCAACCCACCGTGTGCTAGGGCTTCAAGCCCTCCCATGTGGGAC 5156
 QY 581 GlnMetTyrluSvalLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 5157 CAGATGTGAGAGTGTGATTTGGCTCAAGCCCAACCTCCATGGCCCAACCCCTGCTA 5216
 QY 601 TyrluArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrlyTyrlu 620
 DB 5217 TACAGACTGGGCGCTGTCAAGATGAATCACCTTCAACCAACCACTCAACCAATACATC 5276
 QY 621 MetThrCysMetSerlyAspLeuGluValAlaThrSerThrTyrluValLeuValGlyGly 640
 DB 5277 ATGACATGATGTGGCCCACTCGAGGTGTGTACAGACCTGGGTGTCTAGTGGCGGC 5336
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrluCysLeuSerThrGlyCysValValIleValGlyArg 660
 DB 5337 GTCTGTGGCTTGTGGCCCGATTTGCTGTCAACAGGTGGGTGTCTAGTGGCGGC 5396
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrluArgGluPhe 680
 DB 5397 GTGCTTGTTCGGGAAGCGGCAATCATCTGACAGAGGAAGTCTTACAGAGATTCC 5456
 QY 681 AspGluMetGluGluCys 686
 DB 5457 GATGAGATGAAAGATGC 5474

RESULT 6
 US-08-440-549-9
 ; Sequence 9, Application US/08440549
 ; Patent No. 6312889
 ; GENERAL INFORMATION:
 ; APPLICANT: Houghton, Michael
 ; APPLICANT: Choo, Qui-Lim
 ; APPLICANT: Kuo, George
 ; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
 ; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Chiron Corporation
 ; STREET: P.O. Box 8097 (Int. Prop. R-440)
 ; CITY: Emeryville
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94662-8097
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/440,549
 ; FILING DATE: 12-MAY-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/910,760
 ; FILING DATE: 07-JUL-1992

```

1  ATTORNEY/AGENT INFORMATION:
2  NAME: Blackburn Esq., Robert P.
3  REGISTRATION NUMBER: 30,447
4  REFERENCE/DOCKET NUMBER: 0101.002
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE: (510) 601-2702
7  TELEFAX: (510) 655-3542
8  INFORMATION FOR SEQ ID NO: 9:
9  SEQUENCE CHARACTERISTICS:
10 LENGTH: 9401 base pairs
11 TYPE: nucleic acid
12 STRANDEDNESS: double
13 TOPOLOGY: linear
14 MOLECULE TYPE: DNA (genomic)
15 FEATURE:
16 NAME/KEY: CDS
17 LOCATION: 342..9374
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: 366
21 OTHER INFORMATION: /note= "This amino acid position
22 OTHER INFORMATION: can also be Arg."
23 FEATURE:
24 NAME/KEY: misc_feature
25 LOCATION: 372
26 OTHER INFORMATION: /note= "This amino acid position
27 OTHER INFORMATION: can also be Thr."
28 FEATURE:
29 NAME/KEY: misc_feature
30 LOCATION: 867
31 OTHER INFORMATION: /note= "This amino acid position
32 OTHER INFORMATION: can also be Thr."
33 FEATURE:
34 NAME/KEY: misc_feature
35 LOCATION: 1341
36 OTHER INFORMATION: /note= "This amino acid position
37 OTHER INFORMATION: can also be Val."
38 FEATURE:
39 NAME/KEY: misc_feature
40 LOCATION: 2148
41 OTHER INFORMATION: /note= "This amino acid position
42 OTHER INFORMATION: can also be Ile."
43 FEATURE:
44 NAME/KEY: misc_feature
45 LOCATION: 2883
46 OTHER INFORMATION: /note= "This amino acid position
47 OTHER INFORMATION: can also be Asn."
48 FEATURE:
49 NAME/KEY: misc_feature
50 LOCATION: 3681
51 OTHER INFORMATION: /note= "This amino acid position
52 OTHER INFORMATION: can also be Ser."
53 FEATURE:
54 NAME/KEY: misc_feature
55 LOCATION: 3690
56 OTHER INFORMATION: /note= "This amino acid position
57 OTHER INFORMATION: can also be Thr."
58 FEATURE:
59 NAME/KEY: misc_feature
60 LOCATION: 4167
61 OTHER INFORMATION: /note= "This amino acid position
62 OTHER INFORMATION: can also be Leu."
63 FEATURE:
64 NAME/KEY: misc_feature
65 LOCATION: 4323
66 OTHER INFORMATION: /note= "This amino acid position
67 OTHER INFORMATION: can also be Val."
68 FEATURE:
69 NAME/KEY: misc_feature
70 LOCATION: 4701
71 OTHER INFORMATION: /note= "This amino acid position
72 OTHER INFORMATION: can also be Tyr."
73 FEATURE:

```

```

NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-549-9

Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
Gaps: 0
US-09-930-591-2 (1-666) x US-08-440-549-9 (1-9401)
1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyValenGlyCysIleIleThr
:::|||||

```


Dh 3417 CTGGCCGCTACAGCGGCTACGCCAGACAAAGGCGCTCTAGGCTCATATACACC 3476
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Dh 3477 AGCTTACCTGGCGGGGCAAAAACCAAGTGAAGGTGAGGTCCAGATTGTGTCAACTGCT 3536
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleHisGlyAla 60
Dh 3537 GCCCAACCTTCCGGCAACGTGATCATATGGGGTGTGTGGACTGTCTACACAGGGGCC 3596
Qy 61 G1ThrArgThrIleAlaSerProLysGlyProValIleGlnMetIleThrAsnValAsp 80
Dh 3597 GGAACTGAGAACATCGGTGACCCCAAGGGTCTGTCTATCCAGATGATACCAATGATGAC 3656
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Dh 3657 CAAGACTGTGGGCTGGCCGCTCCGAGTGAAGCGCTCATTTGACACCTGTGACTTGC 3716
Qy 101 GlySerSerAspLeuIleValThrArgHisAlaAspValIleProValArgArgArg 120
Dh 3717 GGGCTCTCGGACCTTTACTGTGTACAGAGGACCGCATGTCTATCCGTCGCGCGGCG 3776
Qy 121 G1AspGlyArgGlySerLeuLeuSerProArgProIleSerIleLeuLysGlySerSer 140
Dh 3777 GGTGATGACAGGGGACGCTGTGTGCCCCGCGGACATTTTACCTTGAAGAGCTCTCG 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Dh 3837 GGGGGTCCGCTGTGTGTGCCCCGCGGGACGCGCTGGGCATTTTAGGGCCGCGTGTGC 3896
Qy 161 ThrArgGlyValAlaValAlaValAspPheIleProValGlnSerLeuGlnThrThrMet 180
Dh 3897 ACCGTGTGAGTGTGCTAAGCGGTGAGCTTTATCCCTGTGAGAACCTAGAGAACATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerIleGlnVal 200
Dh 3957 AGGTCCCCGGTGTTCAGATTAACCTCTCTCCACAGATAGTCCCGAGACTTCCAGAGTG 4016
Qy 201 AlaHisLeuHisAlaProThrGlySerGlySerThrIleValProAlaAlaValAla 220
Dh 4017 GCTCACTCATGTCTCCACAGAGGCGGCAAAAAGCACAAAGTCCCGGCTGCATATGCA 4076
Qy 221 AlaGlnGlyIleIleValIleValIleAsnProSerValAlaAlaThrMetGlyPheGly 240
Dh 4077 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTGTGTGTGCAACATGGGGCTTTGCT 4136
Qy 241 AlaIleThrMetSerIleValHisGlyIleAspProAsnIleArgThrGlyValAlaGThrIle 260
Dh 4137 GCTTACATGTCCAAAGCTCATGGATGTGATCTTAACATCAGAACCGGGGTGAGAACAT 4196
Qy 261 ThrThrGlySerProIleThrIleThrIleThrGlyLysPheLeuAlaAspGlyGlyCys 280
Dh 4197 ACCACTGGAGCCCATCATCTACTCCACTACGGCAAGTTCCTTGCCGAGCGGGGTGC 4256
Qy 281 SerGlyValAlaIleValIleIleIleCysAspGlnCysHisSerThrAspAlaThrSer 300
Dh 4257 TCGGGGGGCGCTTATGATCATATTAATTGTGACAGATGCCACTCCAGATGCCAATCC 4316
Qy 301 IleLeuGlyIleIleGlyThrValIleAspGlnAlaGlnThrAlaGlyAlaGleuThrVal 320
Dh 4317 ATTTTGGGCAATGGCACTGTCTTGAACAGAGAGCTCGGGGGGAGAACATGTGTGTG 4376
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGlnVal 340
Dh 4377 CTGGCCACCGCCACCTCCGGGCTCCGTACTGTGTGCCCCATCCCAATCGAGAGGTT 4436
Qy 341 AlaLeuSerThrThrGlyGlnIleProPheTrpGlyValAlaIleProLeuGlnAlaIle 360
Dh 4437 GCTCTGTCACACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTATC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerIleValCysAspGlnLeuAlaIle 380
Dh 4497 AAGGGGGGAGACATCTCATCTTGTCTGCAATTCAAAGAAAGTGGCAGAACTCGCGCA 4556

Qy 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTrpIleArgGlyLeuAspValSerVal 400
Dh 4557 AAGTGTGTGCTATGGGCATCATATGCGGTGCTACTACAGCGGTCTTGAAGTGTCCGT 4616
Qy 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Dh 4617 ATCCGACACGCGCGCATGT 4676
Qy 421 G1AspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Dh 4677 GGGAATTCGACTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4736
Qy 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
Dh 4737 CTGACCTTACCTTACCATTTAGACATAGCATAGCATAGCATAGCATAGCATAGCATAGCAT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTrpArgPheValAlaProGly 480
Dh 4797 CAACGTGCGGGCAGAGCTGTGCAAGGGGAGCCAGGATCTTACAGATTGTGTGCAACGGGG 4856
Qy 481 GluArgProSerGlyMetPheAspSerValLeuGlnGlyCysTrpAspAlaGlyCys 500
Dh 4857 GAGCGCCCTCCGACATGTGTGACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4916
Qy 501 AlaTrpTrpGlyLeuThrProAlaGlnThrThrValArgLeuArgAlaTrpMetAsnThr 520
Dh 4917 GCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheTrpGlnGlyValPheTrpGlyLeu 540
Dh 4977 CCGGGGCTTCCCGTGTGCGCAGACCACTTGAATTTTGGAGGGCGCTTTTACAGGCGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerIleGlnAsnLeuProTrp 560
Dh 5037 ACTCATATGATGATGCCACTTGTATCCAGACAAACAGAGTGGAGAACTTCTCTTAC 5096
Qy 561 LeuValAlaTrpGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Dh 5097 CTGTCAGCGTACCAACGACCGGTGTGCTGAGGGCTCAAGCCCTCCCATGTGTGGAG 5156
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Dh 5157 CAGATGTGAAGTGTGTGATTCGCTCAAGCCCACTCCATGTGGCCAAACCCCTGTCTA 5216
Qy 601 TyrArgLeuGlyValAlaGlnAsnGlnValThrLeuThrHisProValThrIleGlyTrpIle 620
Dh 5217 TACAGACTGGCGCTGTGTGAATGAATCACTTACGACGACCCAGCTCAACAAATATATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGlnValIleThrSerThrTrpValLeuValGlyGly 640
Dh 5277 ATGACATGATGTGCGGCCGACCTCGAGGTGTGTACAGACACCTGGGTGTCTGTGGCGGC 5336
Qy 641 ValLeuAlaAlaLeuAlaIleValIleValIleValIleValIleValIleValIleVal 660
Dh 5337 GTCCTGCTGCTTGTGGCCGGGTATGCTGTGTCAAGGCTGCTGTGTGTGTGTGTGTGTGT 5396
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGlnValLeuTrpArgGlnPhe 680
Dh 5397 GTGCTTGTGTGCGGAAGCGGCAATCATACGACAGGAAATCTTACACGAGATTC 5456
Qy 681 AspGlnMetGlnGlnCys 686
Dh 5457 GATGAGATGGAAGATGC 5474

RESULT 7:

US-08-823-895A-25

Sequence 25, Application US/08823895A

Patent No. 643159

GENERAL INFORMATION:

APPLICANT: Kevin P. Anderson

TITLE OF INVENTION: Compositions And Methods For

TREATMENT OF Hepatitis C Virus-Associated Diseases

NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 66 E. Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,895A
FILING DATE: March 17, 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/453,085
FILING DATE: May 30, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/945,289
FILING DATE: September 10, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 810-1454
INFORMATION FOR SEO ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 9401
TYPE: Nucleic
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO
US-08-823-895A-25

Alignment Scores:
Pred. No.: 0
Score: 3574.00
Percent Similarity: 99.42%
Best Local Similarity: 97.96%
Query Match: 98.78%
Length: 9401
Matches: 672
Conservative: 10
Mismatches: 4
Indels: 0
Gaps: 0

US-09-930-591-2 (1-686) x US-08-823-895A-25 (1-9401)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3417 CTGGCGCCATCATCGGCTACGCCACGACGACAGGGGCTCTTAAGGTGATATACCC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 3477 AGCTTAACGTGGCCGGGCAAAAACCAAGTGAAGGTGAGTCCAGATTGTGTCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaIleThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60
DB 3537 GCCCAACCTTCCTGGCAACGTCATGAGGTGCTGACGTCTTACCAAGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3597 GGAAGAGAGACATCGCGTCAACCCAAAGGTCTCTGTCATCCAGATGTATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3657 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTACCCGCTCATTTACACCCCTGCACTTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 3717 GGCTCCTCGGACCTTTACTGCTGATCAAGAGCACGCCGATGTATTCCTCGTGGCGGGCGG 3776

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer 140
DB 3777 GGTGATAGAGAGGGGAGCGCTGTGTCGCCCGGCCCATTTCTCTTAAGAGGCTCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3837 GGGGGTCCGCTGTTGTGCCCCGGGGGACCGCTGGGCAATTTAGAGCCGGGTGTGC 3896
QY 161 ThrArgGlyValAlaAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3897 ACCCTGAGTGGCTTAAGCGGTGAGCTTTATCCCTGGAGAGACTGAGACACCATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3957 AGTCCCGCGGTGTCAAGGATACCTCTCCACAGTAGTCCCAAGAGCTTCCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIleTyrAla 220
DB 4017 GCTCACCTTCATGCTCCACAGGACGGGCAAAAGACCAAGTCCCGGCTGCATATGCA 4076
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4077 GCTCAGGCTATTAAGTGTCTAATCAACCCCTGTGTGTCGCAACTGGGCTTTGGT 4136
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 4137 GCTTACATGTCCAAAGCTCATGATGATCGATCTTAACATCAGACCGGGGTGAACAATT 4196
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 4197 ACCACTGGAGGCCCATCAACGTAATCCACTCAAGCAAGTTCCTTCCGACGGCGGTGC 4256
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 4257 TCGGGGGGGCTTATGACATTAATATTGTGAGAGTGCACACTCCAGATGCCACATCC 4316
QY 301 IleLeuGlyTyrGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4317 ATCTTGGGCAATCGGCACTGTCTTGAACCAAGACGACGCGGGCGAGACTGTTGTG 4376
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4377 CTGCGCACCGCACCCCTCGGGCTCGTCACTGTCGCCCATCCAACTCGAGGAGTTC 4436
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 4437 GCTCTGTCACCAACCGAGAGATCCCTTTTACGGCAAGCTATCCCTCGAAGTAAATC 4496
QY 361 LysGlyGlyArgHisIleuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 4497 AAGGGGGGAGACATCTCATCTTCTGTCAATTCAAAGAGAGTGCACAGACTCCGCCGA 4556
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4557 AAGCTGTATCCATTTGGGCATCAATGCCGTGCTTACACCGCGCTTGAACGTCCGTC 4616
QY 401 IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
DB 4617 ATCCGACAGAGCGCGATGTTGTGTCGTGGCAACGATCCCTCATGACCGGCTATACC 4676
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4677 GGCACCTTCACTCGGATGATGACATCAATCTGTCTCAACCAAGACAGTCAATTTACG 4736
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4737 CTGACCTTCACTTCAACATTTAGACAAATACCGCTCCCAAGATGTGTCTCCCGCACT 4796
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleLysArgPheValAlaProGly 480
DB 4797 CAACGTGGGGGAGAGATCGGACAGGGGGAACCGAGCATTCACAGATTGTGGCACCGGGG 4856
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500

Query Match:	98.67%	Indels:	0
DB:	3	Gaps:	0
US-09-930-591-2	(1-686)	x US-08-444-818-08	(1-8316)

QY	1	MetAlaProIleThrAlaValAlGlnIleThrArgIleuLeuGlyCysIleIleThr	20
Db	2734	CTGGCGCCCAATCAGCGGCTACGCCACAGACAAAGGGCCCTCTAAGGTGCAATACCC	2793
QY	21	SeuIeThrGlyArgAspIysAsnGlnValGlnGlyIuValGlnIleValSerThrAla	40
Db	2794	AGCCTAACGTGGCCCGGACAAACCAAGATGAGGGGTGAGTCCAGATTGTGTCAACTGCT	2853
QY	41	AlaGlnIhrPheLeuAlaIhTrCysIleIeAsnGlyValCysTrpThrValIyrHisGlyAla	60
Db	2854	GCCCAACCTTCTCGGCAACGTGCATCATGGGGGTGTGTGACTGTCTACACAGGGCC	2913
QY	61	GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetIyrThAsnValAsp	80
Db	2914	GGAAACGAGAACCATCGGTGCACCCAGAGGTCTCTGTCAACAGATGTATACCAATGTAGAC	2973
QY	81	GlnAspIeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	2974	CAAAACCTTGTGGGCTGGCCCGCTCCGCAAGTGAAGCCGTATTGACACCCGTGACCTTGC	3033
QY	101	GlySerSerAspIeuIyrLeuValIThrArgHisAlaAspValIleProValArgArgArg	120
Db	3034	GGGCTCTCGGACCTTTACCTGTGTACACGAGGACGCGCATGTATATCCGTGCGCGGGGG	3093
QY	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerIyrLeuIysGlySerSer	140
Db	3094	GGTATATGACAGGGGACAGCTGTGTGTGCGCCCGGCCAATTTCTTCTTAAGAGGCTCTG	3153
QY	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	3154	GGGGGTCCGCTGTGTGTGTGCCCCGCGGACCCCGTGGCATATTATGAGGCGCGGTGTGC	3213
QY	161	ThrArgGlyValAlaIyAlaValAlaAspPheIleProValGlnSerLeuGlnIhThrMet	180
Db	3214	ACCGGTGAGTGGCTTAAGCGGTGGACTTTATCCCTGTGGAAACCTTAGAACACACATGG	3273
QY	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerIyrGlnVal	200
Db	3274	AGGTCCCGGAGTTACAGGATATACCTCTCCACACAGTAAAGTCCCAAGCTTCCAGGTG	3333
QY	201	AlaHisIleuHisAlaIProThrGlySerGlyIysSerThrIyIysValProAlaAlaTyrAla	220
Db	3334	GCTACCTCCATGTCTCCACAGGACGCGCAAAAGACCAAGAGTCCCGCTGCATATAGCA	3393
QY	221	AlaGlnGlyIyrIyIysValIleuValIleuAsnProSerValAlaIleThrMetGlyPheGly	240
Db	3394	GCTAAGGGCTTAAAGTGTCTAGTACACCCCTCTGTGCTGCACACTGGGCTTTGGT	3453
QY	241	AlaIyrMetSerIyIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	3454	GCTTACATGTCCAAAGGCTCATGGGATGATCTTAACATCAGACCGGGGTGAGAACAATT	3513
QY	261	ThrThrGlySerProIleThrIyrSerThrIyrGlyIyIysPheLeuAlaAspGlyGlyCys	280
Db	3514	ACCACTGGCAAGCCCATCATCAGTACTCCACCTTAAGGCAAGTTCCTTGGCGAGGGGGGTC	3573
QY	281	SerGlyGlyAlaIyrAspIleIleIleCysAspGlnCysHisSerThrAspAlaThrSer	300
Db	3574	TGGGGGGGGCGTTATGACATTAATATTGTGTATGACAGTGCACCTCACGGGATGCCACATCC	3633
QY	301	IleIeuGlyIleGlyIhrValIleuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal	320
Db	3634	ATCTTGGGCATCGGCACATGCTCTTGACCAACACAGACCTGGGGGGCAGACTGTGTGTG	3693
QY	321	IeuAlaIhrAlaIhrProProGlyIysSerAlaIhrValProHisProAsnIleGlnGlnVal	340
Db	3694	CTGGCAACCGCACCCCTCGGGGCTCGGTCACTGGCCCATATCCCAACATGAGAGAGTTT	3753

Oy	341	AlaleuseerThrThngUlleProPheThyGlyValaIleProLeuGlnAlaIle	360
Db	3754	GCTGTCTCCACCAACGGAGAGATCCCTTTTACGGCAAGGCTACCCCTCGAAGTATC	3813
Oy	361	LysGlyGlyAlaGhiIleuIlePheCysHisSerLysLysLysCysAspGlnIleuAlaIle	380
Db	3814	AAGGGGGGAGACATCTCATCTTCTGTCTCAATCAAGAAAGATGCGACGAATCGGCCGA	3872
Oy	381	LysIleuValAlaIleuGlyValAspAlaValAlaIleThyTrpArgGlyIleuAspValSerVal	400
Db	3874	AAGGTGTTCGATTGGGCGATCAAGCCGTGGCTACTACCGCGGCTTGAAGTGCCTC	3933
Oy	401	IleProThSerSerGlyAspValValValValAlaThrAspAlaIleuMetThrGlyPheThr	420
Db	3934	ATCCGACACAGCGCGCAGATGTTGTGTGTGGCAACCATGCTTCATACCGGGCTATAC	3993
Oy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrIleValIleAspPheSer	440
Db	3994	GGCAGCTTCGACTGGATGATGACTGCAATCGTGTCAACCCAGACAGTGCATTCAGC	4053
Oy	441	LeuAspProThrPheThrIleGlyIleThrIlePheIleuProGlnAspAlaValSerArgThr	460
Db	4054	CTTAACTTCATCTTACCATTTGAACAATCAAGCTCCCAAGAAAGCTGTCTCCGCACT	4113
Oy	461	GlnArgArgGlyArgThnGlyArgGlyLysProGlyIleThyArgPheValAlaProGly	480
Db	4114	CAAGCTGGGGGACAGACTGGCAGGGGAGGAGCAGCATCTACAGATTTGTGTGCACGGGG	4172
Oy	481	GlnArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTrpAspAlaGlyCys	500
Db	4174	GAGGCGCCCTCCGCGCATGTTGCACTCGTCCTCTGTGTGATCTGACGACGAGCTGT	4233
Oy	501	AlaIleThyTrpGlyIleuThrProAlaGlyIleThrThrAlaArgLeuArgAlaIleMetAsnThr	520
Db	4234	GCTTGGTATGACTCAGCCCGCCGAGACTACAGTTAGCTACAGAGGCTACATGAACCC	4293
Oy	521	ProGlyIleuProValCysGlnAspHisIleuGlnIleThrProGlnGlyValPheThrGlyIleu	540
Db	4294	CCGGGGCTTCCCGTGTGCGACAGACCATTTGAAATTTGGAGGGGGCTTTACAGGCTC	4353
Oy	541	ThrHisIleAspAlaHisPheIleuSerGlnThrLysGlnSerGlyIleAsnIleuProTrp	560
Db	4354	ACTCATATAGATGCGCCACTTTCTATCCACAGCAAGACGAGTGGGAGAACTTCTTAC	4413
Oy	561	LeuValAlaIleArgIleAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp	580
Db	4414	CTGGTATGATCCAGACCACTGTGGCTATAGGCTCAAGCCCTCCCAATCGTGGAAC	4477
Oy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrIleuHisGlyProThrProIleuLeu	600
Db	4474	CAGATGGGAAGTGTTTGATTCCGCTCAAGCCCACTCCATGGGGCCAAACCCCTGCTA	4533
Oy	601	ThyArgLeuGlyAlaValGlnAsnGlyValAlaThrIleuThrHisProValThrLysTrpIle	620
Db	4534	TACAGACTGGGCGCTGTTCAGAAAGAAATCACTGTACGCAACCAAGTACCAAAATACATC	4593
Oy	621	MetThrCysMetSerAlaAspIleuGlyValValIleThrSerThrTrpValIleuValGlyGly	640
Db	4594	ATGCAATGCAATGGCGCGACCTGAGAGTCTCAAGACACCTGGAGTCTGTGTGGCGGC	4653
Oy	641	ValIleuAlaAlaIleuAlaAlaIleCysIleuSerThnGlyCysValValIleValGlyArg	660
Db	4654	GTCCTGGCTGTTGGCGCGGTATTGGCTGTCAACAGGCTCGTGTGATATGTGGGACGG	4713
Oy	661	IleValIleuSerGlyLysProAlaIleIleProAspArgGlyValLeuTrpArgGlyPhe	680
Db	4714	GTCCTCTTGTTCGGGGAAGCGGCAATCATCTGACAGGGAAGTCTCTTACCGAGAGTTC	4773
Oy	681	AspGlnMetGlnGlyCys686	
Db	4774	GATAGATGGAAGAGTGC4791	

```

US-08-444-818-122
; Sequence 122, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Ruter, William J.
; TITLE OF INVENTION: NANBY Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110,002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-444-818-122

Alignment Scores:
Pred. NO.: 0 Length: 9185
Score: 3570.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5
Query Match: 98.67% Indels: 0
Gaps: 3

US-09-930-591-2 (1-686) x US-08-444-818-122 (1-9185)

```

```

QY 101 GlySerSerAspLeuThrValThrArgHisAlaAspValIleProValArgArg 120
DB 3695 GGCTCTCGGACCTTACTTGCTGACGAGGACCGCCGATGTCATCCGTCGCGCGG 3754
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuIleGlySer 140
DB 3755 GGTGATAGAGGGGAGCTGCTGTCGCGCCGCAATTTCTTACTTGAAAGGCTCTCG 3814
QY 141 GlyGlyProLeuLeuCyseProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3815 GAGGAGTCGCTGTGTGCCCCCGGAGCAGCGGTGGCATTTTAAAGGCCCGGTGTGC 3874
QY 161 ThrArgGlyValAlaAlaValAlaAspPheIleProValGlySerLeuGlyThrThrMet 180
DB 3875 ACCCGTGAAGTGAAGCGGTGAGCTTATCCCTGTGAGAGCACTGAACAACATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlySerTyrGlnVal 200
DB 3935 AGGTCCCGGTGTTCACGATTAATCTCTCCACAGTATGTCGCCAGAGCTTCCAGGTG 3994
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3995 GCTCACTTCATAGCTCCACAGGACGCGCAAAAGCAAGCTCCGCGCTGCATATGCA 4054
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4055 GCTCAAGGCTATAGGTGCTAGTACCAACCCCTGTTGCTGCAACACTGGGCTTTGGT 4114
QY 241 AlaTyrMetSerLysValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 4115 GCTTACATGTCAGAGGCTCATGGATGCAATCTTACATAGAACCGGGGTGAGAACATTT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 4175 ACCACTGGACAGCCCATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 4234
QY 281 SerGlyGlyAlaTyrAspIleIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
DB 4235 TCGGGGGGCGGTATATCAATTAATTTGTGAGAGTGCACCTCCAGATGCAATCC 4294
QY 301 IleLeuGlyTyrIleGlyThrValLeuAspGlnAlaGlyThrAlaGlyAlaArgLeuThrVal 320
DB 4295 ATCTTGAGCATTCGCACTGCTCTTGAACAGAGAGACTGCGGGGCGAGACTGTTGGT 4354
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlyIleVal 340
DB 4355 CTGGCCACGCGCACCCCTCGGGCTCGTCACTGTCGCCCATCCCAATCCGAGAGGTT 4414
QY 341 AlaLeuSerThrThrGlyGlyIleProPheTyrGlyLysAlaIleProLeuGlnAlaIle 360
DB 4415 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGGAAGGCTATCCCTCCGAGATATC 4474
QY 361 LysGlyGlyLysArgHisLeuIlePheCysHisSerLysLysCysAspGlyLeuAlaAla 380
DB 4475 AAGGGGGGAGACATCTCATCTTCTGTCACTTCAAGAGAGAGCGACAGAACTCGCGCA 4534
QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4535 AAGTGTGTCATTTGGGCATCAATCGCTGCTCACTACCGCGGTCTTGAACGTGCTCGTC 4594
QY 401 IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
DB 4595 ATCCCGACCAAGCGGCGAGTGTGTGTGTGGCAACGATGCCCTCATAGACCGGCTATAC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4655 GCGCACTTCACTCGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4714
QY 441 LeuAspProThrPheThrIleGlyThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4715 CTTGACCTCACTTCACTTGAAGACATATGACATGATGATGATGATGATGATGATGATG 4774
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyLysLeuThrPheValAlaProGly 480

```


DB 4775 CAACGTCGGGCGAGCTGGCAGGGGGAGCCAGGCACTTACAGATTGTGGACACCGGG 4834
QY 481 GUAAGProSerGlyMetPheAspSerValLeuGlyCysTyrAPAlaGlyCys 500
DB 4835 GAGCGCCCTCCGGGCACTGTCAGTCCGTCCGTCTGTAGTGTGACGAGGCTGT 4894
QY 501 AlaTPTyGlyLeuThrProAlaGlyThrThValArgLeuArgAlaTyrMetAsnThr 520
DB 4895 GCTTGGATAGAGCTCAAGCCCGGAGACATACAGATTAGCTACAGAGGTACATGAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
DB 4955 CCGGGGCTTCCCGTGGCCAGACCACTTGAATTTGGAGGGCCGCTTTTACAGGGCTC 5014
QY 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrArgGlnSerGlyValAsnLeuProTyr 560
DB 5015 ACTCATATAGATGCCCACTTCTATCCCAACAAAGAGTGGGAGAACCTTCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrPasp 580
DB 5075 CTGGTAGCGTACCAAGCCAGCTGTGGCTTAGGGCTCAAGCCCTCCCATCGTGGAC 5134
QY 581 GlnMetTyrGlyCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGTGAAGTGTGATTGCTCCACACCCCACTCCATGGGCCCAACACCCCTGCTA 5194
QY 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 5195 TACAGACTGGGCGCTGTTCAAGATGAATACCCCTTACGCAACCCATCACAATACATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTyrValLeuValGlyGly 640
DB 5255 ATGACATGATGTCGGCCGACCTGAGGTGCTGCACGAGCACCTGGGGTCTGTTGGGGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 5315 GTCTGTGCTGCTTGGCGCGGTATGCTGTCAACAGGCTGCGTCAATAGTGGGAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 5375 GTCTGTGCTGCGGAGGCGGCAATCATCTGACAGGAGTCTTACCGAGGTTC 5434
QY 681 AspGluMetGluGlyCys 686
DB 5435 GATGATGTGAAGATGTC 5452
RESULT 11
US-08-444-818-123/C
Sequence 123, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESSES:
ADDRESSES: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 123:
SEQUENCE CHARACTERISTICS:
LENGTH: 9185 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ANTI-SENSE: YES
US-08-444-818-123

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0	3570.00	9185	671	10	5	0	
Percent Similarity:	99.27%						
Best Local Similarity:	97.81%						
Query Match:	98.67%						
DB:	3						

US-09-930-591-2 (1-686) x US-08-444-818-123 (1-9185)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 5791 CTGGCCCATACAGCGCTGACCCCGCAGACAAAGGGGCTCTTAGGGTGCATTAACACC 5732
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
DB 5731 AGCTTACTGGCCGGACAAACCAAGTGAAGGTGAGGTCCAGTTGTGTAACCTGCT 5672
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
DB 5671 GCCCAACTTCTCTGGCAACGTGCATCATGGGGTGTCTGGACGTCTTACCAACGGGGC 5612
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 5611 GGAACGAGGACCATCGCTCACCCCAAGGGTCTGTATCATCAATGATGAGAC 5552
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 5551 CAAGACTTGTGGGCTGGCCGCTCCGCAAGTACCGCTCAATGACACCTGCACTTGC 5492
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 5491 GGCTCTCGGACCTTACCTGTGTACAGAGGACCGCGCATCTATCCCTGCCCGCGG 5432
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 5431 GGTGATAGAGGGGAGCTGTGTCGCCCGGCCCATTTCTTCACTTGAAGGCTCCTCG 5372
QY 141 GlyGlyProLeuLeuLysCysProAlaGlyHisAlaValGlyTyrPheArgAlaAlaValCys 160
DB 5371 GGGGCTCGCTGTGTGGCCCGGGGACGCGGTGGGAGATTTTGAAGGCCCGGGTGC 5312
QY 161 ThrArgGlyValAlaLysValAlaValAspPheIleProValGlnSerLeuGlnThrThrMet 180
DB 5311 ACCCGTGAAGTGGCTTAGGCGGTGACCTTATCCCTGTGAGAACTTAGACACACCATG 5252
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 5251 AGTCCCGCGTGTCAAGGATACCTCTCCACCAAGTAGTGCCCAAGAGCTTCCAGGAGT 5192
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrThrValProAlaAlaTyrAla 220
DB 5191 GCTCACTTCATGCTCCCAAGGACGGGCAAAAGCAACAAAGTCCGGCTGCATATGCA 5132


```

QY 221 AlaGInGlyTyrIlyValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 5131 GCTCAGGGGTAAAGGGTGAAGTACTCAACCCCTGTTGCTGCAACACTGGGCTTGGT 5072
QY 241 AlaTyrMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 5071 GCTTACATGTCACAGGCTCATGGGATCGATCTTACATCAAGACCGGGGGTGAACAACATT 5012
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIysPheLeuAlaAspGlyIys 280
DB 5011 ACCACTGGAGCCCATCAAGTACTCCACTACGGCAAGTCTCTGGCCGAGCGGGGTGC 4952
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlnCysHisSerThrAspAlaThrSer 300
DB 4951 TCGGGGGGCGGTATGACATTAATTTGACAGAGTCCACTCCAGGATGCCAATCC 4892
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4891 ATCTGGGCAATCGGCATCTGCTTGACCAAGCAAGACTCGGGGGGGAAGCTGGTTGTG 4832
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4831 CTGGCCACCGGCACCTCCGGGGCTCCGTCACGTGCTCCATCCCAACATCGAGAGGTT 4772
QY 341 AlaLeuSerThrThrThrGlyGluIleProPheThrGlyIysAlaIleProLeuGluAlaIle 360
DB 4771 GCTCTGTCACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAGATATC 4712
QY 361 LysGlyIysIleArgHisLeuIlePheCysHisSerIysIysCysAspGluLeuAlaIle 380
DB 4711 AAGGGGGGAGACATCTCATCTTCTGTCATTCANAGAAAGATCGACGAAGTCCGCCGA 4652
QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4651 AAGCTGGTCGATGGGCATCAATGCCGTGCTTACACCGCGGTCTTGAGCTGTCCGTC 4592
QY 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 4591 ATCCCGACGAGCGGCGATGTTGCTGCTGCGCAACCGATGCCCTCAAGACCGGTATACC 4532
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4531 GGGCCTTCGACTCGGTGATGACTGCAATACGTGTGCAACCCAGACAGTGCATTCAGC 4472
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4471 CTGACCTTCACTTCAACCATTAAGACATCAAGTCCCTGCTGATGATGACGCGGCGCT 4412
QY 461 GlnArgArgGlyArgThrGlyArgGlyIysProGlyIleTyrArgPheValAlaProGly 480
DB 4411 CACGCTCGGGGAGACAGTGCAGGGGAGAACCAAGCATCTACAGATTTGTGGCACCGGGG 4352
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 4351 GAGCGCCCTCCGCGCATGTCGACTCGTCCGCTCTGCTGATGATGACGCGGCGCT 4292
QY 501 AlaTyrGlyArgLeuThrProAlaGluThrValArgLeuAspAlaTyrMetAsnThr 520
DB 4291 GCTTGGTATGAGCTACGCGCCGACGACTACAGTTAGGCTACGAGCTTACATGACACC 4232
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
DB 4231 CCGGGGCTTCCGCTGTCAGGACCATCTTGAATTTTGGAGGGGCTCTTACAGGCTTC 4172
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIysGlnSerGlyGluAsnLeuProTyr 560
DB 4171 ACTCATATGATGCCACACTTCTATCCGACAAAGACAGAGTGGGAGAAACCTTCTTAC 4112
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerThrAsp 580
DB 4111 CTGGTAGCGTACCAAGCCACCGTGTGCGCTTAGGGCTCAAGCCCTCCCTCCATCGGGAGC 4052
QY 581 GlnMetTyrIysCysLeuIleArgLeuIysProThrLeuHisGlyProThrProLeuLeu 600

```

```

DB 4051 CAGATGGAGATGTTGATTGCTCGCTCAAGCCCAACCCCTCAATGGGCCCAACACCCCTGCTA 3992
QY 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrIleTyrIle 620
DB 3991 TACAGATCGGGGCGGTTCAGATGAATATCAACCTTGACGACCCAGTCAACCAATACATC 3932
QY 621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTyrValLeuValGlyGly 640
DB 3931 ATGACATGATGATCGGCCGACCTGAGAGGTGTCACAGACACCTGGGTGCTGTTGGCGGC 3872
QY 641 ValLeuAlaIleLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 3871 GTCCTGCTGCTTGGCGCGCTTATGCTGTCACAGCGTGGGTGATGATGGGCAAG 3812
QY 661 IleValLeuSerGlyIysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 3811 GTCGCTTGTCCGGGAGCGGCAATCATCTGACAGGAGTCTTACCGAAGATTCT 3752
QY 681 AspGluMetGluGluCys 686
DB 3751 GATGAGATGGAAGAGTGC 3734

RESULT 12
US-09-881-239-2
; Sequence 2, Application US/09881239
; Patent No. 6630298
; GENERAL INFORMATION:
; APPLICANT: CHIRN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESK, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COIT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
; FILE REFERENCE: 2302-16073 / P16073.003
; CURRENT FILING DATE: US/09/881,239
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
; NAME/KEY: CDS
; LOCATION: (1)..(2058)
US-09-881-239-2

Alignment Scores:
Pred. No.: 0 Length: 2058
Score: 3565.00 Matches: 671
Percent Similarity: 99.13% Conservative: 9
Best Local Similarity: 97.81% Mismatches: 6
Query Match: 98.54% Indels: 0
DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-09-881-239-2 (1-2058)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1 ATGGCGCCCATCAACGGGTGACGCCCAACAGAGGGGCTCTTACGGGTGCATTAATACCC 60
QY 21 SerLeuThrGlyArgAspIysAsnGlnValGlyGluValGlnIleValSerThrAla 40
DB 21 AGCTTAATGCGCGGACAAACCAACATGAGAGGTAGGTCAAGATTGTCAACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 121 GCCCAACCTTCCCTGGCAACGTGATCAATGGGGGTGCTGACTGTCTACCAACGGGGCC 180

```

QY 61 GlyThrArgThrIleAlaSerProGlySerGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 181 GGAAACGAGGACCAATCGGGTCAACCAAGGCTCTGTCTCCAGATGATATACCAATGTAGAC 240
QY 81 GluAspLeuValGlyTTPProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 241 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGTACCGATCATTTGACACCTGTGACTTGC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGCTCCTCGGACCTTATCTGTGTACGAGGACGCGCATTCATTCGCGCGCGCGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlySerSer 140
DB 361 GGTGATGACAGGGGACCTGTGTCCCGCGCCATTTCTTCACTTGAAGAGGCTCTCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 421 GGGGGTCCGCTGTGTGCCCCGCGGGACGCGTGGCATTTAGGGCCGCGGTGTC 480
QY 161 ThrArgGlyValAlaAlaValAlaAspPheIleProValGlnSerLeuGlnThrThrMet 180
DB 481 ACCGTGAGATGGCTAAGCGGTGAGTCTTATCCCTGTGAGAACTAGAGACACCATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTTCCTCGGTGTCCAGATTAACCTCTTCCACCACTAGTGGCCCGACGCTTCCAGGTG 600
QY 201 AlaHisLeuHisAlaProThrArgIleSerGlySerGlySerThrIleValProAlaAlaTyrAla 220
DB 601 GCTCACTCTCAATGCTCCACAGGAGCGCAAAAGCCAAAGTCCGCGCTGCAATAGCA 660
QY 221 AlaGlnGlyTyrIleValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGGCTAATAGGTGCTAGTACTCAACCCCTCTCTCTGCAACACTGGGCTTTGCT 720
QY 241 AlaTyrMetSerIleValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCCAAAGGCTCATGGGATGATCTTAACATCAGACCGGGGGGAGAAACAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIlePheLeuAlaAspGlyGlyCys 280
DB 781 ACCACTGGCAGCCCCCAATCAAGTACTCCACTACGCGCAAGTCTTGGCGAGCGGGTGC 840
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
DB 841 TCGGGGGGCGCTTATGACATATATTTGTGACGAGGCCACTCCACGAGATGCCCATTC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
DB 901 ATCTTGGGCAATGGGCACTGTCTTGAACAAGACACTCGGGGGGAGAGACTGTGTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGlnVal 340
DB 961 CTGCGCACCGCACCCCTCGGGCTCGGTCACTGTGCCCATCCCAACATCGAGAGGTT 1020
QY 341 AlaLeuSerThrThrGlyGlnIleProPheTyrGlyValAlaIleProLeuGlnAlaIle 360
DB 1021 GCTCTCTCCACCAACCGAGAGATCCCTTTTACGGCAGGCTATCCCTCGAGATATC 1080
QY 361 LysGlyGlyValArgHisLeuIlePheCysHisSerIleValCysAspGlnLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCTCTGTCAATTAAGAAAGAGGCGAGCGAAGCTCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1141 AAGCTGCTGCATTTGGCATCAATGCGTGGCTTACTACCGCGCTTGAACGCTGCTGC 1200
QY 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCGCCCATCGGCGATGTGTGCTGTGTGCAACGAGATCCCTCATGACCGGCTATAC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

DB 1261 GGGACTTGACCTCGGTGATAGCTGCAATACGTGTGTCAACCAAGTCAATTCAGC 1320
QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCTTACCTTCACTTGAACATATACGCTTCCCAAGATCTGTCTCCGCACT 1380
QY 461 GlnArgArgGlyValArgThrGlyValArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTGGGGGAGACCTGCGAGGGGAGACCGGATCATCAATGATTTGTGGACCGGG 1440
QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGlnCysTyrThrAspAlaGlyCys 500
DB 1441 GAGCGCCCTCGGCAATGTTCAGCTGTCCGTCTCTGTGAGGTGATAGCAGAGCTGT 1500
QY 501 AlaTyrTyrGlnLeuThrProAlaGlnThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGTATGAGCTTACGCGCCCGCGAGACTACATTAAGCTTACGACGATACAAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheTyrGlnGlyValPheThrGlyLeu 540
DB 1561 CCGGGGCTTCCGCTGTGCGCAGGACCATCTTGAATTTGGAGGGGCTCTTTACAGGCTTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleValGlnSerGlyGlnAsnLeuProTyr 560
DB 1621 ACTCATATGATGCTCCCATTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
DB 1681 CTGTATGCTTACCAACCCACCGTGTGCTTGGGCTCAAGCCCTTCCCATCTGTGGAC 1740
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1741 CAGATGTGAAGTGTGATTTGGCTCAAGCCCACTCCATGGGCAACACCCCTGCTA 1800
QY 601 TyrArgLeuGlyAlaValGlnAsnGlnValThrLeuThrHisProValThrIleValIle 620
DB 1801 TACAGACTGGGGCTGTGTGATGATGAATACCTGTGACGACCCAGTCCCAATATCATC 1860
QY 621 MetThrCysMetSerLysAspLeuGlnValIleThrSerThrTyrValLeuValGlyGly 640
DB 1861 ATGACATGATGTGCGCCCACTGAGGTGTGTCACAGCACTGGGTGCTCTGTGGCGG 1920
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValAlaIleValGlyArg 660
DB 1921 GTCTGTGCTGTGGCCCGCTATTCCTGTCAACAGCTGCTGTATAGTGGCAGG 1980
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGlnValLeuTyrArgGlnPhe 680
DB 1981 GTCTGTGTGCGGGAGCGGCAATCATACCTGACAGGAGAGTCTTACCGAGAGTTC 2040
QY 681 AspGlnMetGlnGlnCys 686
DB 2041 GATGAGATGAGAAAGTGC 2058
RESULT 13
US-09-881-654-1
; Sequence 1, Application US/09881654
; Patent No. 6532601
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESKA, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COLT, Doris
; TITLE OF INVENTION: IMMUNOSSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / P17039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811

PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/280,867
PRIOR FILING DATE: 2001-04-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 1
LENGTH: 2058
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence:
OTHER INFORMATION: representative NS3/4a conformational antigen
NAME/KEY: CDS
LOCATION: (1) .. (2058)
US-09-881-654-1

Alignment Scores:
Pred. No.: 0 Length: 2058
Score: 3565.00 Matches: 671
Percent Similarity: 99.13% Conservative: 9
Best Local Similarity: 97.81% Mismatches: 6
Query Match: 98.54% Indels: 0
DB: Gaps: 0

US-09-930-591-2 (1-686) x US-09-881-654-1 (1-2058)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1 ATGGCGCCATCAGCGGGTACGCCAGACAGAGGGGCTCTTGGGTGATATATACC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 61 AGCTTAATGCGCCGGGCAAAAACCAAGTGGAGGGTGAAGTCCAGATTGTCAACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleHisGlyVal 60
DB 121 GCCCAACCTTCCCTGGGCAACGTGCATCAATGGGGTGTCTGGACTGTCTACCAAGGGGCC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 181 GGAGCGAGACATCGGCTCACCCAGGGTCTCTTCATCCAGATGTATCCAAATGTAGAC 240
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 241 CAAGACCTGTGGGCTGGCCGCTCCGCAAGGTAGCGATCATTGACACCCCTGACCTTGC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGCTCCCTCGGACCTTTACCTGTGTACAGAGGACCGCATGTCTCCCTGGCGCGGGCGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 361 GGTATATGACAGGGGACGCTCTCTGCGCCGCGCCATTCTTCTTGAAGAGCTCTCTCG 420
QY 141 GlyIleProLeuLeuCysProIleGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 421 GGGGGTCCGCTGTGTGCGCGCGGGGACGCGTGGGACATATTAAAGGCGCGGCTGTGC 480
QY 161 ThrArgGlyValAlaIleValAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 481 ACCCGTGAAGTGGCTTAAGCGGTGAATTTTCCCTGTGGAGAACTTAGAGAACCAATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTCCCGGTGTTCAAGGATTAATCTCTTCCACCAAGTAGTCCCGCAGAGCTTCAAGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 601 GCTACACCTCATGCTCCACAGGACAGCGGCAAAAGACCAAGAGTCCGGCTGCATATGCA 660
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGCTATTAAGTGTAGTACTCAACCCCTCTGTGTGCAACACTGGGGCTTTGGT 720

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCCAGGCTCATGGATCATCTTCAATCATCAGGACCGGGGTGAGAACAAAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 781 ACCACATGGACGCCCCATCAAGTACTCACCTACGGGCAAGTTCCTTCCAGCGGGGTGC 840
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
DB 841 TCGGGGGGGCTTATATACATATATTTGTGACGAGTGCACCTCCAGGATGCAATCC 960
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValAlaArgLeuThrVal 320
DB 901 ATCTGGGCAATTGGCACTGCTCTTACCAAGACAGACTGCGGGCGGAGACTGTTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 961 CTTCGACACCGCCACCCCTCGGGCTCGTCACTGTGCCCATCCCAACATCGAGAGATT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGlnAlaIle 360
DB 1021 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGCTATCCCTCCAGATATATC 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGlyLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCATCTTCTGTCAATCAAGAGAGTGCAGACAACTCGCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAlaAlaValAlaTyrThrArgGlyLeuAspValSerVal 400
DB 1141 AACCTGGTGCATTTGGGCATCAATGCGGCTTACCTACCGGGTGTGACGTTGCGTCC 1200
QY 401 IleProThrSerGlyAspValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCGCCCATTCGGCGATGTGTCTGTGTGCAACCGAGTCCCTCATGACCGGCTATAG 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 1261 GCGGACTTGCACCTCGGTATAGACTCAATACGTGTGTCCAGCCAGACAGTCTTTCAGC 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCCATCACTTCAACATGAGACATACGCTCCCAATGCTGTCTCCGACT 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTGCGGGGACGAGCTGGCAGGGGAGAACGAGCATCTACGATTTGTGGACCGGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1441 GAGGCGCCCTCCGGCATGTTCGACTGCTCGTCTGTGATGCTATGACGAGGCTGT 1500
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetLeuThr 520
DB 1501 GCTTGTATATAGCTTCAGCGCGCGGACCTACAGTTAAGGTACAGAGGTACATGAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
DB 1561 CCGGGGCTTCCGTGTGCCAGACATCTTGAATTTTGGGAGGGCGCTTTTACAGGCTTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysLeuLeuProTyr 560
DB 1621 ACTCATATATAGTCCCACTTCTATCCCAAGACAAACAGAGTGGGAGAACTTCTTACC 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB 1681 CTGTATGCGTACCAAGGACCGGTGTGCGCTAAGGAGCTCCATCCCATGTGGGAGAC 1740
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyLysProThrProLeuLeu 600
DB 1741 CAGATGTGAAGTGTGATTGCTTCCGCTCAGCCACCTCATGTGGCCAAACACCCCTGTCTA 1800

QY 601 TyrArgLeuGluValAlaGlnAsnGluValThrLeuThrHisProValThrIleValGlyTyrTle 620
DB 1801 TACAGACTGGCCCTGTTCCAGAAATGAATCACTTCAGCAGCACCAGTCAACCAATATCATC 1860
QY 621 MetThrCysMetSerLysAspLeuGluValThrSerThrTrpValLeuValGlyGly 640
DB 1861 ATACATGACATGTCGGCCGACCTGAGAGTGTACAGAGCCTGGGGTGTGTTGGGGGC 1920
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 1921 GTCTGTGCTCTTGTGGCCGGTATGTGCTTCACAGGCTGCGTGTCAATAGTGGGAGG 1980
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 1981 GTGCTCTTGTCCGGGAAGCCGGCAATCATCTGACAGGAGGTCTCTACCGAGATTTC 2040
QY 681 AspGluMetGluGluCys 686
DB 2041 GATGAGATGAGAGATGTC 2058
RESULT 14
US-08-444-818-137
; Sequence 137, Application US/08444818
; Patent No. 615087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110,002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8987 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8985
; US-08-444-818-137
Alignment Scores:
Pred. No.: 0 Length: 8987
Score: 3565.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5
Query Match: 98.54% Indels: 0

DB: 3 Gaps: 0
US-09-930-591-2 (1-686) x US-08-444-818-137 (1-8987)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3076 CTGGCCCATCATCAGCGGCGTACGCGCCAGCAGACAGAGGGGCTCTTACGGTGCATATCAC 3135
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 3136 AGCTTAACCTGGCCGGGACAAACCAAGTGAAGGGTGAAGTCCAGATTGTCAACTGCT 3195
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleGlyAla 60
DB 3196 GCCCAACCTTCCTGCGCAACGTGCATCAATGGGGTGTGCTGCACTCTCAACCGGGGCG 3255
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGluMetTyrTrpAsnValAsp 80
DB 3256 GGAACGAGAGCAATCCGTCACCCCAAGGGTCTGTATCCAGTATACCAATGTAGAC 3315
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3316 CAAGACTTGTGGGCTGGCCGCTCCGCAAGGTAGCCGCTCATATGACACCTGCACTTGC 3375
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 3376 GGCTCTCGGACTTATACCTGGTCAAGAGGACGCGCATGTCACTCCGTGGCGCGCGG 3435
QY 121 GlyAspGlyArgGlySerLeuLeuSerProAlaGlyProIleSerTyrLeuLysGlySerSer 140
DB 3436 GGTGATAGAGAGGGGAGCGCTGTGCGCCGGCCCATTTCTTCACTTGAAAGGCTCTCG 3495
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgHisAlaValCys 160
DB 3496 GGGGATCCCTGTGTGTGCCCCGGGGGACGCGGTGGCATTTTGGGCGCGGCTGTGC 3555
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGlnThrMet 180
DB 3556 ACCCGTGAAGTGGCTAAGCGGGTGAAGCTTATCCCTGTGAAGACTTACAGCAACCAAG 3615
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3616 AGGTCCCGCGTTCACGATTAATCTCTCCACAGGTAGTGGCCGAGGCTTCAGG 3675
QY 201 AlaIleLeuHisAlaProThrGlySerGlyLysSerThrValProAlaIleTyrAla 220
DB 3676 GCTCACCTTCATGCTCCACAGGCAAGGCAAAAGCAACCAAGGCTCCGGCTGCATATGCA 3735
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 3736 GCTCAGGGCTATTAAGTGTATGATCAACCCCTGTGTGCTGCACACTGGGGCTTTGGT 3795
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 3796 GCTTACATGTCCAAAGGCTCATGGGATCGATCTTAACATCAGAGCCGGGTGAGAACAT 3855
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 3856 ACCACTGGAGGCCCATCAAGTCACTCCACTTCAGGAAGTTCCTGGCCACGGCGGGTGC 3915
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 3916 TCGGGGGGCGCTTATATCAATTAATTGTGACAGTGCACCTCCACGATGCCACATCC 3975
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 3976 ATCTGGGATCGGACGTGCTTGCACAGCAGAGACTCGGGGGGAGAGCTGGTTGGG 4035
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4036 CTTCGACACCGCACCCCTCCGGGCTCCGTCATGTGCCCATTCACCAATCAGAGGAGTT 4095
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyAlaIleProLeuGlnAlaIle 360

```

Db      4096 GCTCTGTCACCAACCGAGAGATCCCTTTTACGGCAGGCTATCCCTCCGAAGTAATC 4155
Qy      361 LysGlyValArgHisLeuIlePheCysHisSerIleValGlyCysAspGluLeuAlaIa 380
Db      4156 AAGGGGGGGAGACATCTCATCTTCTGTCATTCAGAAAGAGAGCGCAGCAATCGCGCA 4215
Qy      381 LysLeuValAlaLeuGlyValAsnAlaValAlaIleTyrTyrArgGlyLeuAspValSerVal 400
Db      4216 AAGCTGCTCCCATTTGGGCATCAATGCCGTGGCCCTACTACCGCGGTCTTGAAGTCCGTC 4275
Qy      401 ILleProThiSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
Db      4276 ATCCCGACCGACGGCGATGTTGTCGTGTGGCAACCATGCTCCCTCATGACCGGCTATACC 4335
Qy      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      4336 GGGCACTTCACATCGGTGATGACTGCATACGTTGTCAACCGACAGCATGCAATTCAGC 4395
Qy      441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      4396 CTTCACCTTCACTTCACCATTTAGACATACACGCTCCCGACGATGCTCTCCCGCACT 4455
Qy      461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyTyrIleTyrArgPheValAlaProGly 480
Db      4456 CAACGTGCGGAGAGACTGCGAGGGGAGAGCAAGCATCAACAGATTGTGGCAGCCGGGG 4515
Qy      481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyCysTyrAspAlaGlyCys 500
Db      4516 GAGGCCCCCTCCGGCATGTCGATCGATCGTCCGTCTGTGATGATGATGACGAGCGCTGT 4575
Qy      501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db      4576 GCTTGGATATAGCTCAACGCCCGGAGACATACAGTTAGGCTACGAGGCTACAGAAACACC 4635
Qy      521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db      4636 CCGGGGCTTCGCCGTGTCAGACCACTTGAAATTTGGAGGGCGCTTTTACAGGCGCTC 4695
Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyValAsnLeuProTyr 560
Db      4696 ACTCATATAGATGCCCATTTCTATCCAGCAAGAGAGAGTGGGGAACCTTCTTAC 4755
Qy      561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrPasp 580
Db      4756 CTGGTAGCGTACCAAGCCACCGGTGTGCTAGGCTCAACGCCCTCCCATGCTGGGAC 4815
Qy      581 GlnMetTyrPlyCysLeuIleArgLeuLysProThrIleuHisGlyProThrProLeuLeu 600
Db      4816 CAGATGTGGAAGTGTGATTCGCTCAAGCCCACTTCATGGGCCCAACCCCTGCTA 4875
Qy      601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      4876 TACGAGCTGGGCGCTGTTCCAGATGAAATCACCTTGACGACCCAGACCAACAAATACATC 4935
Qy      621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db      4936 ATGACATGCAATGTGGCGACCTGAGAGTGTGTACGACACCTGGGAGTCTGTTGGGCGC 4995
Qy      641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      4996 GTCTGCTGCTGTTGGCGCGCTATTTGCTGTCAACAGGCTGCGGTCAATAGTGGGAGG 5055
Qy      661 ILleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db      5056 GTCTCTTGTCCGGGAGCGGCAATACATACATGACGAGGAATCTCTACCGAGAGTTC 5115
Qy      681 AspGluMetGluGluCys 686
Db      5116 GATGATGAGAGAGTGC 5133

```

RESULT 15
US-08-444-818-53

```

/ Sequence 53, Application US/08444818
/ Patent No. 6150087
/ GENERAL INFORMATION:
/ APPLICANT: Chien, David Y.
/ TITLE OF INVENTION: NANOV Diagnostics and Vaccines
/ NUMBER OF SEQUENCES: 777
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSER: Chiron Corporation
/ STREET: 4560 Horton Street
/ CITY: Emeryville
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94608-2916
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/444,818
/ FILING DATE:
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/403,590
/ FILING DATE: 14-MAR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Harbin, Alisa A.
/ REGISTRATION NUMBER: 33,895
/ REFERENCE/DOCKET NUMBER: 0110.002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (508)359-3876
/ TELEFAX: (508)359-3885
/ INFORMATION FOR SEQ ID NO: 53:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5360 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..5360
/ US-08-444-818-53

Alignment Scores:
Pred. No.: 0 Length: 5360
Score: 3555.00 Matches: 668
Percent Similarity: 99.428 Conservative: 13
Best Local Similarity: 97.524 Mismatches: 4
Query Match: 98.264 Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-53 (1-5360)
Qy      1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      930 CTGGCGCCCATCAACGGCGTACGCGCCAGACAGCAAGGGGCGCTCTAGGGGTGCAATACACC 989
Qy      21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db      990 AGCTTAACCTGGCGGAGCAAAACCAAGTGAAGGTGAGAGTCAAGATGTGTCACTGCT 1049
Qy      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db      1050 GCCCAAACTTCTCTGCAACGTGCATCATATGGGGTGTCTGACCTCTACCAACGGGGCC 1109
Qy      61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      1110 GGAACGAGGACATCGCTCACCAAGGGGTCTGTCATCATGATGATACATGTAGAC 1169
Qy      81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100

```

Db 1170 CAAGACCTTGTGGCGTGGCCGCTCCGCAAGTAGCCGCTATTGACACCCCTGCATTCG 1229
 Qy 101 G1ySerSeraspLeuThyValThrArgHisAlaAspValIleProValArgArgArg 120
 Db 1230 GGGTCTCTGGACCTTACTGTGTGACAGAGGACGCCGATGTCTATCCCGTGGCCGGCG 1289
 Qy 121 G1yAspG1yArgG1ySerLeuLeuSerProArgProIleSerTyrluYsg1ySerSer 140
 Db 1290 GGTGATAGCAGGGGCGAGCTGTGTGTGCCCGGCCCATTTCTTACTTGAAAGGCTCTCG 1349
 Qy 141 G1yG1yProLeuLeuCyPProAlaG1yHisAlaValG1yIlePheArgAlaAlaValCys 160
 Db 1350 GGGGGGTCCGCTGTGTGTGCCCGGGGCGACGCGTGCGCATATTAGGGCGCGGTGTGC 1409
 Qy 161 ThrArgG1yValAlaValAlaValAspPheIleProValG1ySerLeuG1yIleThrMet 180
 Db 1410 ACCCGTGAAGTGTGATAGGGCGGTGACTTATCTCTGTGAGAACTTAGAGACAACTATG 1469
 Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProG1yIleVal 200
 Db 1470 AGGTCCCGGTGTGACGATTAATCTCTCTCCACAGTATGCCCCGAGCTTCAGGCTG 1529
 Qy 201 AlaHisLeuHisAlaProThrG1ySerG1yLysSerThryValProAlaAlaIlyrAla 220
 Db 1530 GCTCAGCTCCATGCTCCACAGGCGAGCGCAAAAGCACAAAGTCCCGCTGCATATGCA 1589
 Qy 221 AlaG1yG1yThryValValLeuValLeuValProSerValAlaAlaThrMetG1yPheG1y 240
 Db 1590 GCTCAGGCTATAGAGGTGATGACTACACCCCTGTGTCTGTGACAACTGGGCTTTGGT 1649
 Qy 241 AlaTyrlMetSerLysAlaHisG1yIleAspProAsnIleArgThryValArgThryIle 260
 Db 1650 GCTTACATGTCCAAAGGTGATGGATGATCTTAACATCAGAACCGGGGAGAGAACATTT 1709
 Qy 261 ThrThryG1ySerProIleThryrThryrThryG1yLysPheLeuAlaAspG1yG1yCys 280
 Db 1710 ACCACTGGCGAGCCCATCAGTACTCCACTACGCGCAAGTCTTGGCGAGCGGGGTGC 1769
 Qy 281 SerG1yG1yAlaThryrAspIleIleIleCysAspG1yLysHisSerThryrAspAlaThrSer 300
 Db 1770 TCGGGGGCGCTTATGACATTAATTTGTGACAGGCGCACTCCAGGATGCCACATCC 1829
 Qy 301 IleLeuG1yIleG1yThryrValLeuAspG1yAlaG1yIleThryrAlaArgLeuThryrVal 320
 Db 1830 ATCTTGGGCGATCGGCACTGTCTTGAACAAAGAGAGCTCGGGGGGAGATCGTGTGTG 1889
 Qy 321 LeuAlaThryrAlaThryrProG1ySerValThryrValProHisProAsnIleG1yG1yVal 340
 Db 1890 CTCGCCACCGCGCACCCCTCGGGGCTCGTCACTGTGCCCATCCCAATCGAGAGGTT 1949
 Qy 341 AlaLeuSerThryrThryrG1yG1yIleProPheThryrG1yValAlaIleProLeuG1yAlaIle 360
 Db 1950 GCTCTGTGCCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTATC 2009
 Qy 361 LysG1yG1yArgHisLeuIlePheCysHisSerLysLysLysCysAspG1yLeuAlaAla 380
 Db 2010 AAGGGGGGAGACATCTCATCTTCTGTCTATTCAAAGAGAGATCGAGCAATCGCCGCA 2069
 Qy 381 LysLeuValAlaLeuG1yValAsnAlaValAlaTyrlThryrG1yLeuAspValSerVal 400
 Db 2070 AAGCTGGTCCATTTGGGCAATGCGGTGCTTACTACGCGGTCTTGAAGTGTCCGTC 2129
 Qy 401 IleProThrSerG1yAspValValValValAlaThryrAspAlaLeuMetThryrG1yPheThr 420
 Db 2130 ATCCCGACGAGCGCGATGTGTGTGTGTGCGCAACGATGCCCTCATGACCGGCTATACC 2189
 Qy 421 G1yAspPheAspSerValIleAspCysAsnThryrCysValThryrG1yThryrValAspPheSer 440
 Db 2190 GGGGACTTCACTCGGATGAGTCAATACGTGTGTCAACCGAGACATCGATTTACG 2249
 Qy 441 LeuAspProThryrPheThryrIleG1yThryrIleThryrLeuProG1yAspAlaValSerArgThryr 460
 Db 2250 CTGACCCCTACCTTCAACATGAGACATCAGCTCCCGAGATGCTGTCTCCGCACT 2309

Qy 461 GlnArgArgG1yArgThryrG1yArgG1yLysProG1yIleTyrlArgPheValAlaProG1y 480
 Db 2310 CAACGTGGGGGAGAGATCTGGACGGGGGAGAGCCAGGATCTACAGATTTGTGGACCGGG 2369
 Qy 481 GlnArgProSerG1yMetPheAspSerSerValLeuCyG1yG1yCysTyrlAspAlaG1yCys 500
 Db 2370 GAGCGCCCTCCCGGCGATGTGTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2429
 Qy 501 AlaThryrG1yLeuThryrProAlaG1yIleThryrThryrValArgLeuAlaThryrMetAsnThr 520
 Db 2430 GCTTGTATAGAGTCAAGCCCGCGAGACTACAGTATGAGCTACAGAGCTACATGAACACC 2489
 Qy 521 ProG1yLeuProValCysG1yAspHisLeuG1yPheThryrG1yG1yValPheThryrG1yLeu 540
 Db 2490 CCGGGGCTTCCCGTGTGCCAGACATCTTGAATTTTGGAGGGGCTCTTACAGGCTTC 2549
 Qy 541 ThrHisIleAspAlaHisPheLeuSerG1yThryrLysG1ySerG1yG1yLysLeuProTyrl 560
 Db 2550 ACTCATATAGATGCCCATTTCTATCCAGACAAAGCAGAGTGGGAGAACTTCTCTTAC 2609
 Qy 561 LeuValAlaTyrlGlnAlaThryrValCysAlaArgAlaG1yAlaProProProSerTyrlAsp 580
 Db 2610 CTGGTATGCTTACCAAGCCACCGTGTGAGGCTCAAGCCCTCCCTCCATCGAGGAC 2869
 Qy 581 GlnMetTyrlPheCysLeuIleArgLeuLysProThryrLeuHisG1yProThryrProLeuLeu 600
 Db 2670 CAGATGTGAAGTGTGTGATGTGCTCCACAGCCCACTCATGGGCAACCCCTGTCTA 2729
 Qy 601 TyrlArgLeuG1yAlaValGlnAsnG1yValThryrLeuThryrHisProValThryrTyrlle 620
 Db 2730 TACAGACTGGGCGCTGTTCAGAAATGAATCAACCTGACGACCACTCATCAACATATCATC 2789
 Qy 621 MetThryrCysMetSerLysAspLeuG1yValValThryrSerThryrThryrValG1yG1y 640
 Db 2790 ATGACATGATGTGTGGCGGACCTGAGGTGTGTCAAGACGCTGGGTGTCTGTGTGCGGC 2849
 Qy 641 ValLeuAlaAlaLeuAlaAlaTyrlCysLeuSerThryrCysValValIleValG1yArg 660
 Db 2850 GTCTGTCTGTCTTGGCCGCGATTTGCTGTCAACAGGCTGTGTGTGTGTGTGTGTGTGT 2909
 Qy 661 IleValLeuSerG1yLysProAlaIleIleProAspArgG1yValLeuTyrlArgG1yPhe 680
 Db 2910 GTGTCTGTGTGCGGAAAGCGGCAATCATACCTGACAGGGAAGTCTTACCGAGAGTTC 2969
 Qy 681 AspG1yMetG1yG1y 685
 Db 2970 GATGAGATGGAAG 2984

Search completed: September 17, 2004, 10:32:31
 Job time : 214 secs

THIS PAGE BLANK (USPTO)